Deep Exclusive Reactions

Lecture 1: Nucleon structure studies with the electromagnetic probe
- Elastic scattering: form factors
- DIS: structure function
- Exclusive reactions: Generalized Parton Distributions

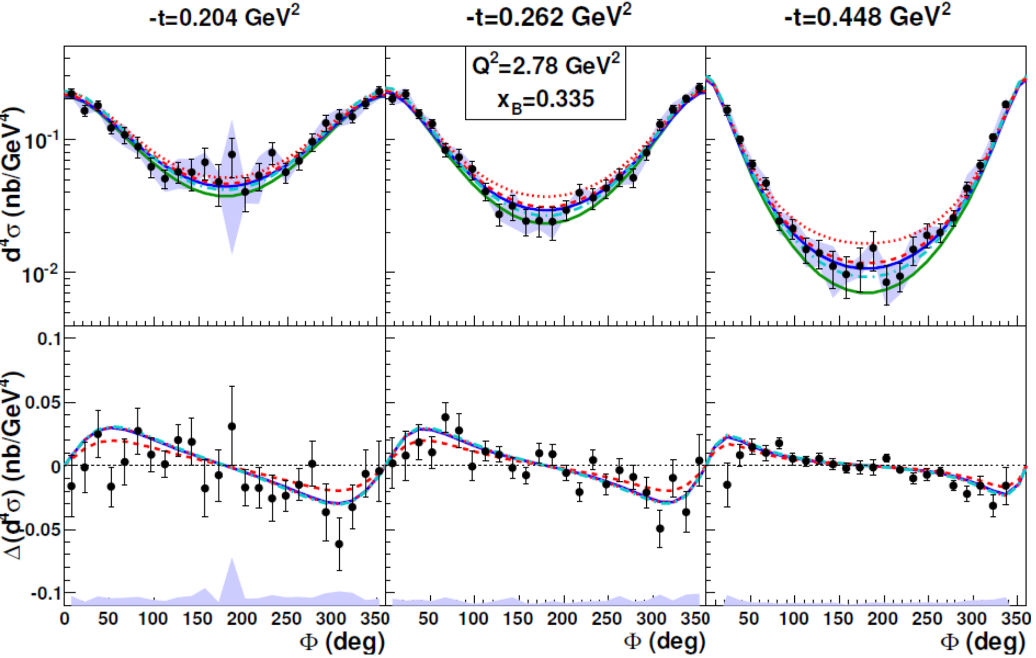
Lecture 2: Deeply Virtual Compton Scattering
- GPDs and DVCS
- DVCS on the proton with JLab@6 GeV
- Extraction of GPDs from data
- Proton tomography and forces in the proton

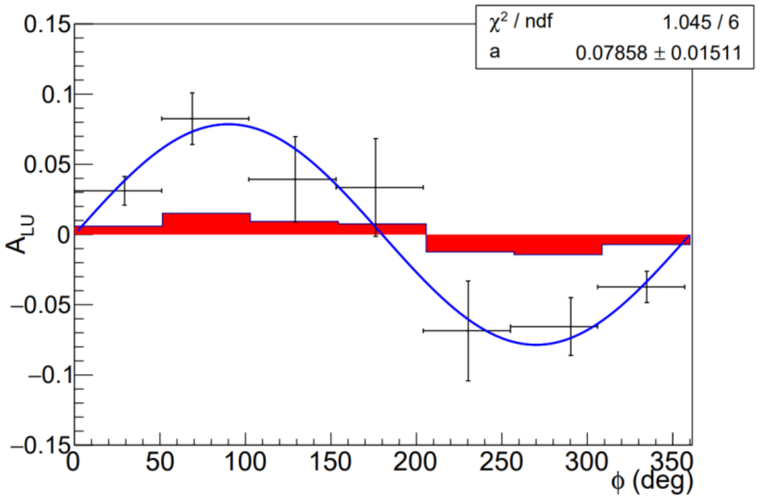
Lecture 3: DVCS and beyond
- DVMP
- New DVCS experiments@12 GeV
- TCS

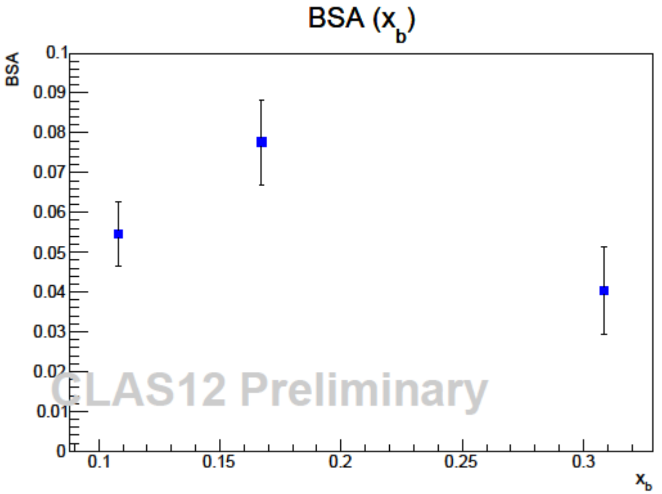
Lecture 4: Perspectives
- Perspectives for a positrons beam
- GPDs at the EIC

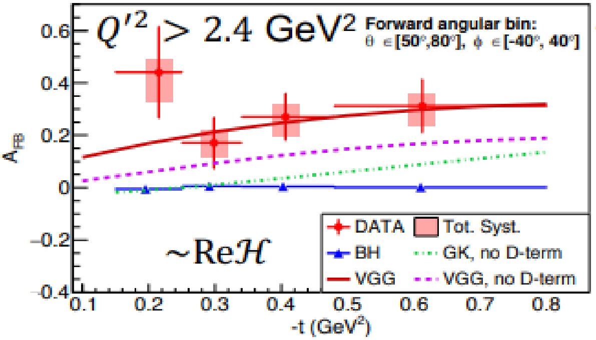
Lecture 5: Tutorial
- Data analysis techniques for exclusive reactions

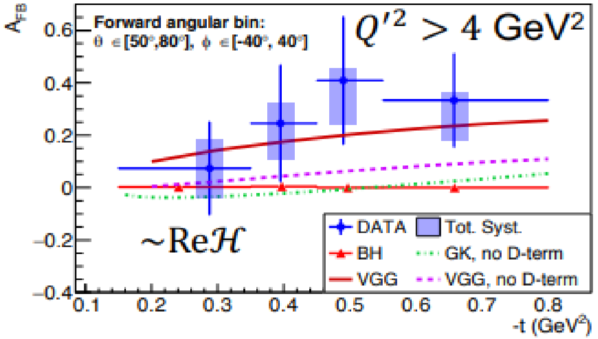
Silvia Niccolai, IJClab Orsay & CLAS Collaboration
HUGS, JLab, June 2023

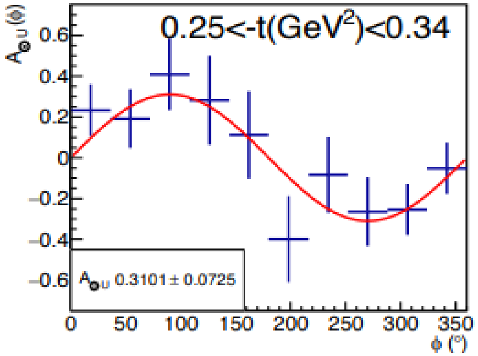


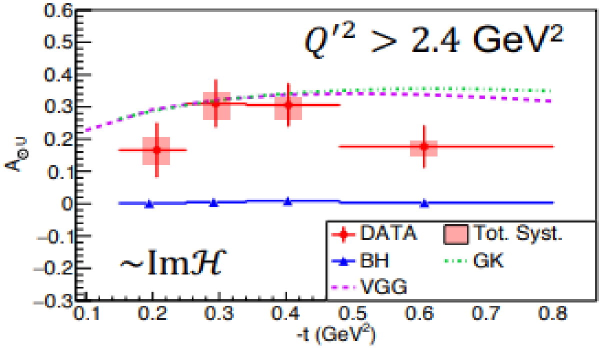










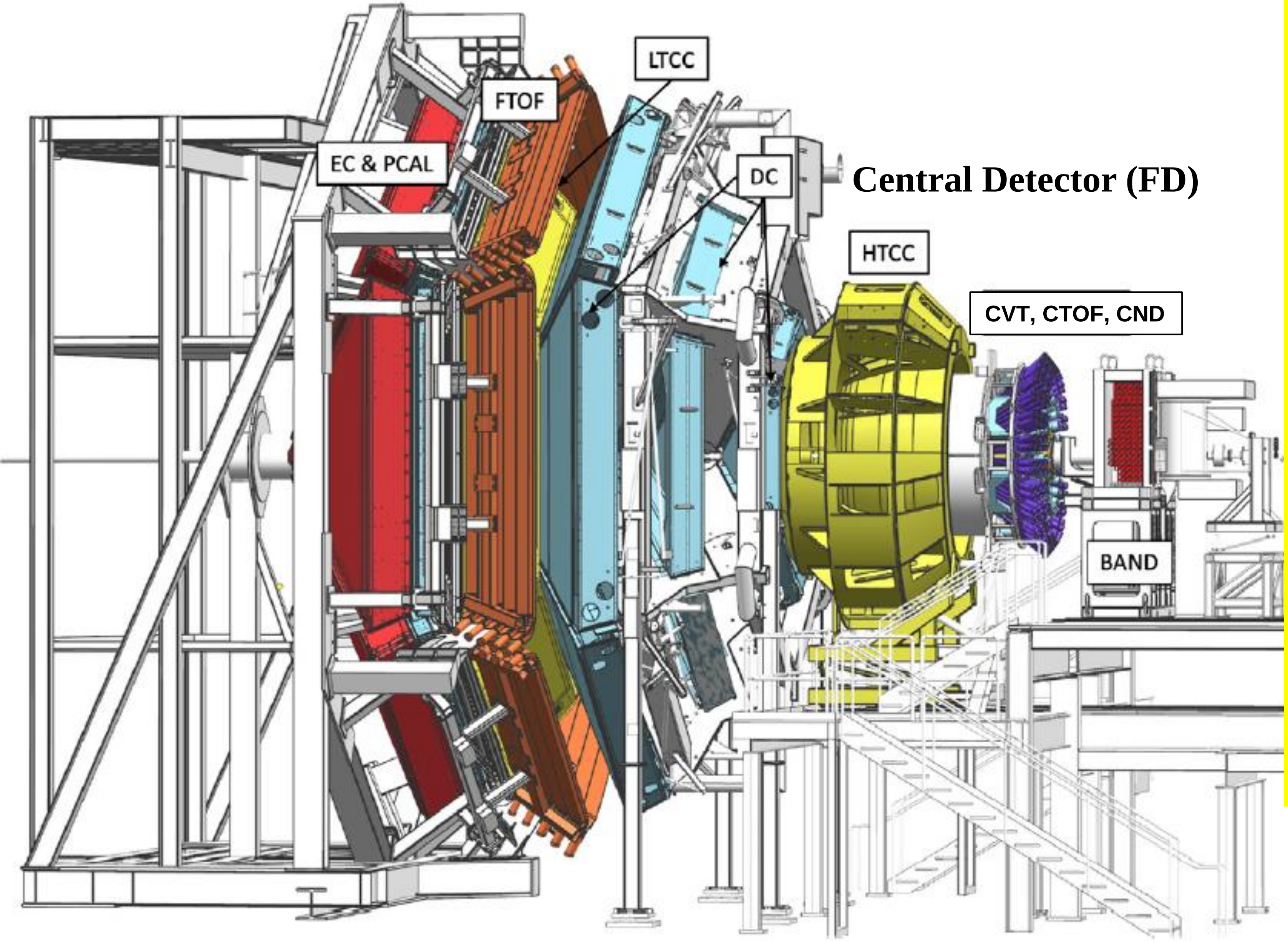


Examples and inspiration taken from:

- *pDVCS cross section analyses (CLAS, Hall A)*
- *nDVCS BSA analysis (CLAS12)*
- *TCS analysis (CLAS12)*
- *ρ^0 cross section analysis (CLAS)*

- Detector example: CLAS12
- Definitions of the observables
- PID (with ML examples)
- Fiducial cuts
- Detection topologies
- Exclusivity cuts
- Background subtraction
- Efficiency/acceptance
- Data-driven efficiency
- Systematic uncertainties

The CLAS12 spectrometer in Hall-B

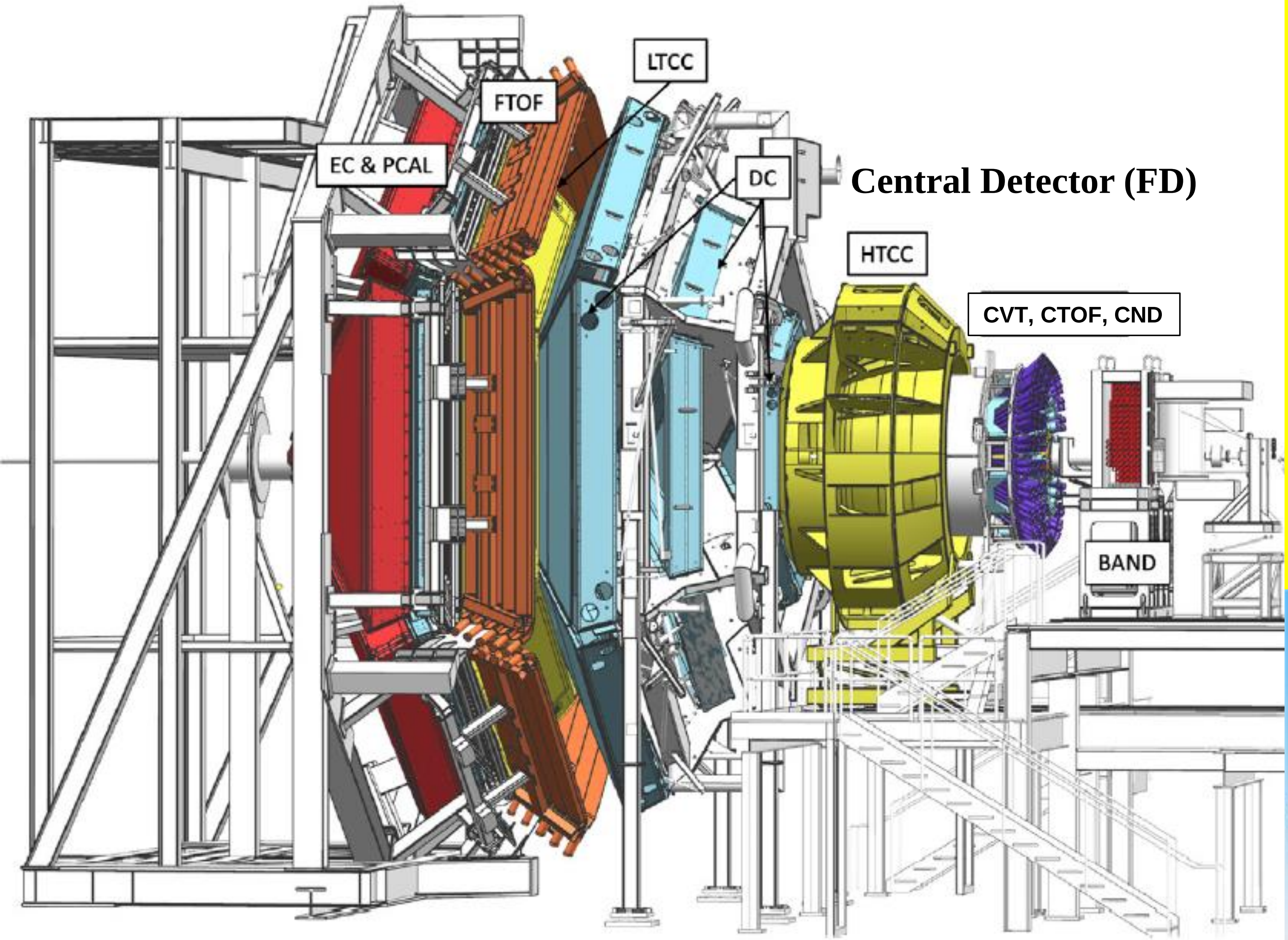


CEBAF Large Acceptance Spectrometer:

- **Efficient detection** of charged and neutral particles over a **large fraction of the full solid angle**
- **Dual-magnet** system:
  - superconducting 6-coils **torus** magnet and detectors at the **forward** polar angles up to 35° and almost-full azimuthal coverage
  - **solenoid** magnet and detectors covering the **central** polar angles from 35° to 125° with full azimuthal coverage
- The Forward Detector is segmented in 6 equally equipped **azimuthal sectors**
- **The target is placed at the center of the Central Detector**

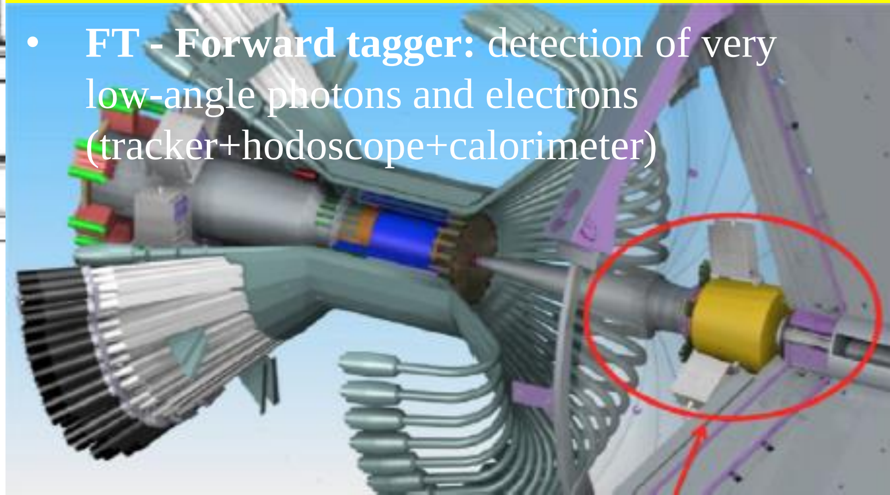
The CLAS12 spectrometer in Hall-B

Forward Detector (FD)



Central Detector (FD)

- **EC&PCAL - Electromagnetic calorimeters:** trigger, PID, detection of neutrals
- **FTOF (CTOF)– Forward (Central) time of flight**: PID
- **DC – Drift chambers**: tracking, determination of the charge and momentum of charged particles
- **HTCC – High-threshold Cherenkov Counter:** trigger, PID
- **CVT - Central Vertex Tracker:** tracking, determination of the charge and momentum of charged particles
- **CND – Central Neutron Detector:** detection of neutrons in the central region



- **FT - Forward tagger:** detection of very low-angle photons and electrons (tracker+hodoscope+calorimeter)

Definition of the observables

$$\frac{d\sigma}{d\phi dx_B dt dQ^2} = \frac{N_{ep\gamma}(\phi, Q^2, x_B, t)}{L_{int} \cdot Acc \ \cdot \Delta Q^2 \cdot \Delta x_B \cdot \Delta t \cdot \Delta\phi \cdot F_{corr}}$$

4-fold differential cross section

$$\text{BSA}\,(\phi, Q^2, x_B, t) = \frac{d\sigma^+\,(\phi,Q^2,x_B,t) - d\sigma^-\,(\phi,Q^2,x_B,t)}{P_B(d\sigma^+\,(\phi,Q^2,x_B,t) + d\sigma^-\,(\phi,Q^2,x_B,t))} = \frac{N^+\,(\phi,Q^2,x_B,t) - N^-\,(\phi,Q^2,x_B,t)}{P_B(N^+\,(\phi,Q^2,x_B,t) + N^-\,(\phi,Q^2,x_B,t))}$$

$N_{ep\gamma}$: yield of the epγ events

$$L_{int} = n_{target} \int_T \frac{dN_e}{dt} dt = n_{target} \frac{Q_{int}}{q_e} = \frac{l_{target} \cdot \rho_{target} \cdot N_A}{M_H} \cdot \frac{Q_{int}}{q_e}$$

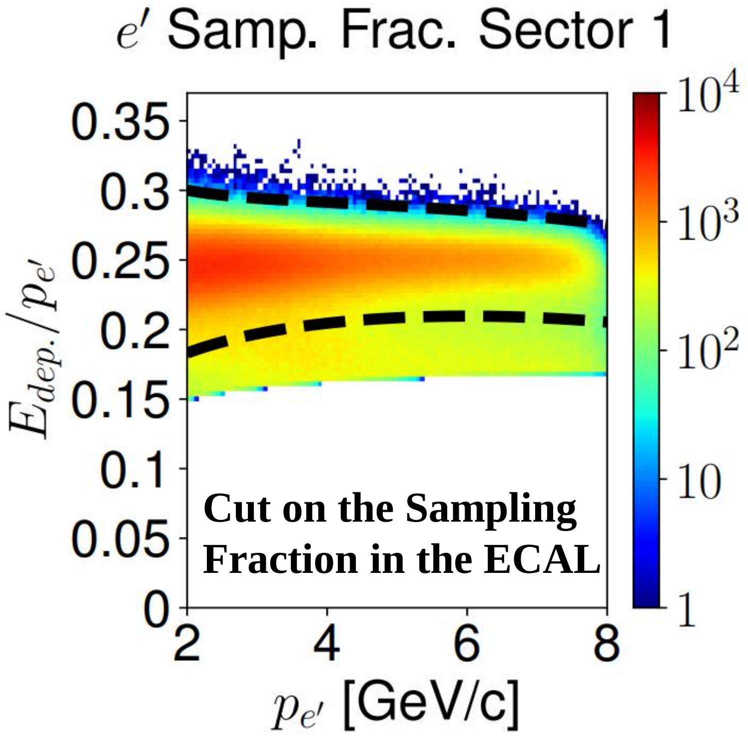
Acc : efficiency*acceptance of the epγ events

$\Delta Q^2 \cdot \Delta x_B \cdot \Delta t \cdot \Delta\phi$: bin hypervolume

P_B : beam polarization, measured during the data taking in dedicated « Moeller » runs

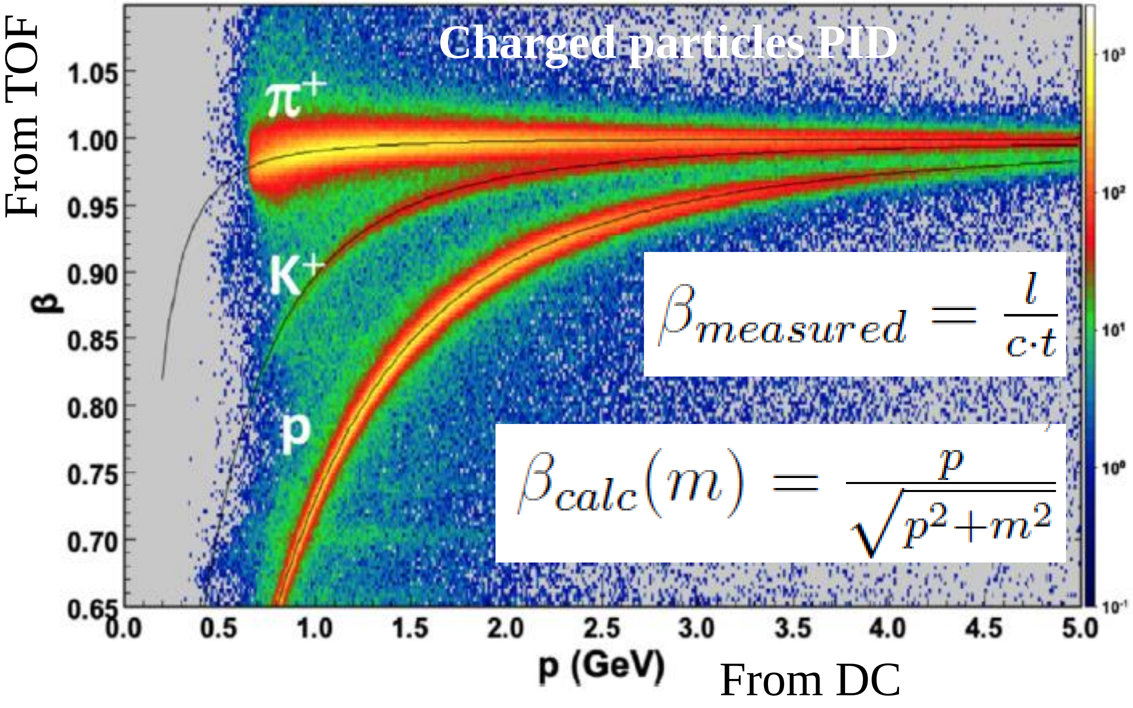
- n_{target}: number of atoms per cm^2 in the target
- dN_e/dt: flux of electrons
- T: duration of the data acquisition
- Q_{int}: integrated livetime-gated charge
- q_e: charge of the electron
- l_{target}: target length
- ρ_{target}: hydrogen target density
- N_A: Avogadro's constant
- M_H: hydrogen molar mass.

PID (Particle IDentification) in CLAS12



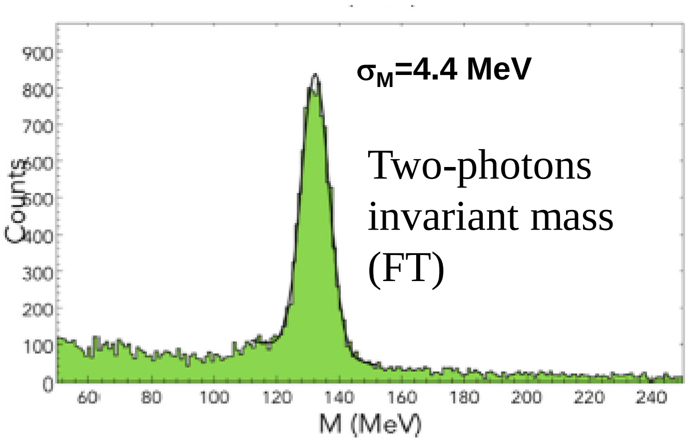
Cut on the Sampling Fraction in the ECAL

- Electrons are detected combining the information of DC (charge-momentum), HTCC (photoelectrons), ECAL
- ECAL: « club sandwich » of lead/scintillator layers; Not all the deposited energy is reconstructed (~1/4 of the electron energy)

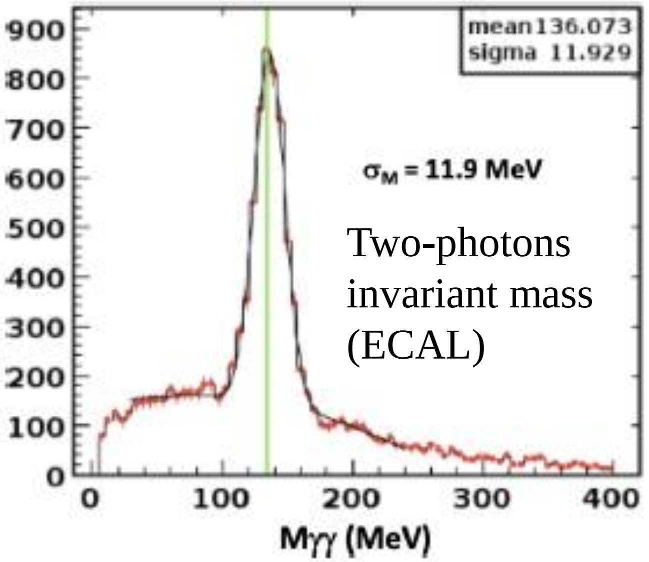


$$\beta_{measured} = \frac{l}{c \cdot t}$$

$$\beta_{calc}(m) = \frac{p}{\sqrt{p^2+m^2}}$$



Two-photons invariant mass (FT)



Two-photons invariant mass (ECAL)

The photons are detected in the ECAL and the FT:

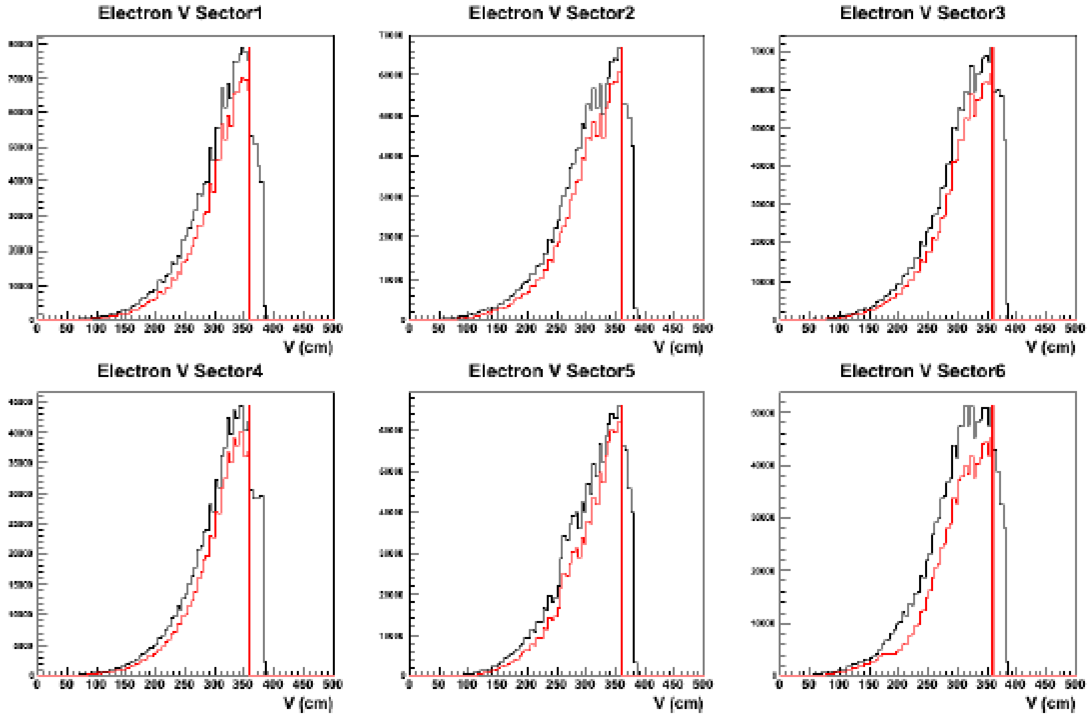
- Calorimeter cluster not geometrically matched to a track→neutral
- β~1 (to remove the neutrons); different cut sizes between ECAL and FT, to account for different resolutions

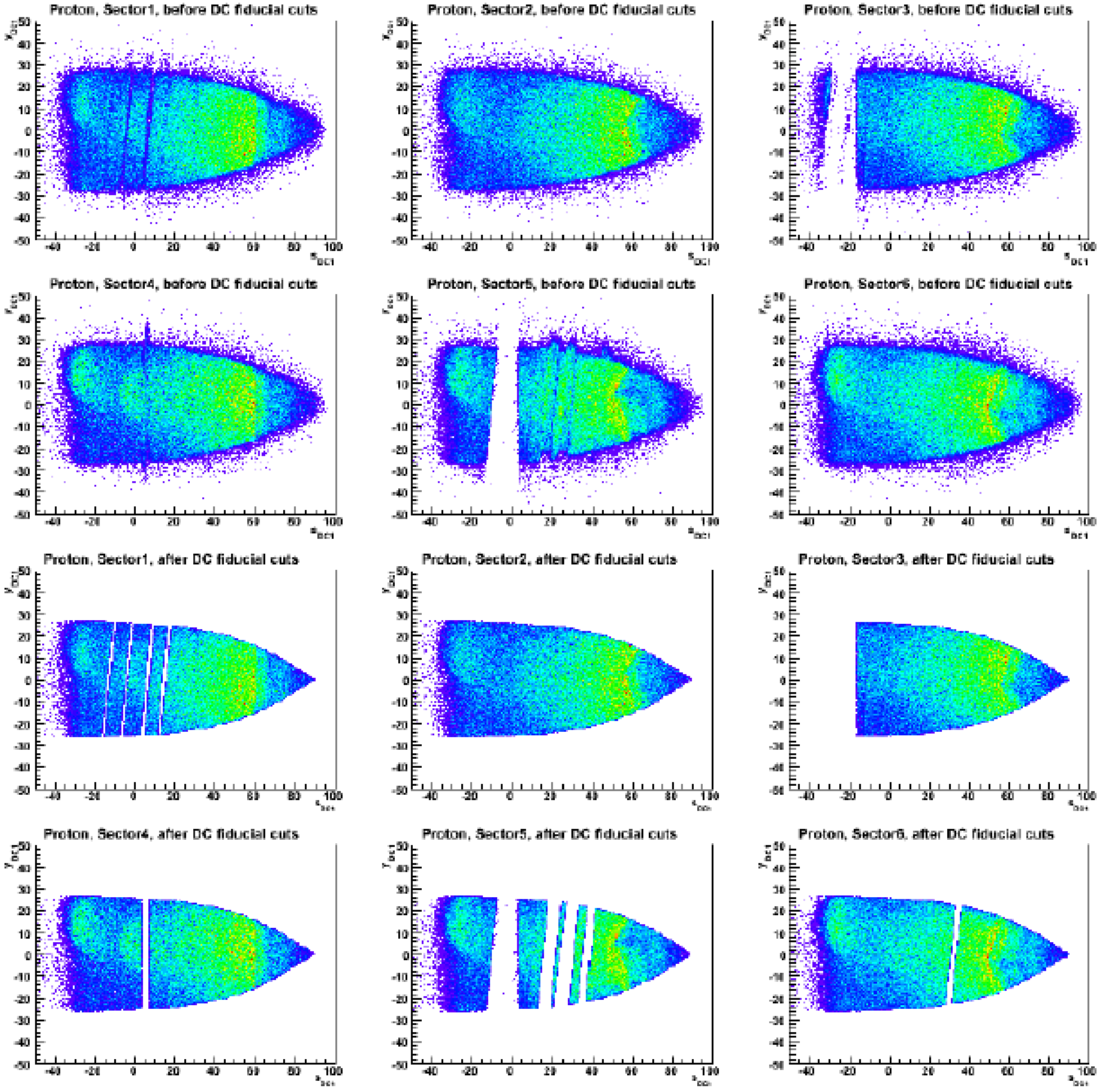
The neutrons are detected in the ECAL and in the CND:

- Cluster not geometrically matched to a track→neutral
- β<1 (to remove the photons)

Fiducial cuts

- Fiducial cuts aim to remove regions of the detectors where the efficiency is not optimal nor well reproduced by simulation (typically, the edges, or « holes »)
- Particularly important in cross-section analyses (where data/MC match is crucial
- Examples from the pDVCS cross-section analysis from CLAS data (6 GeV):
    - DC fiducial cuts for protons
    - ECAL fiducial cuts for electrons, V « view »



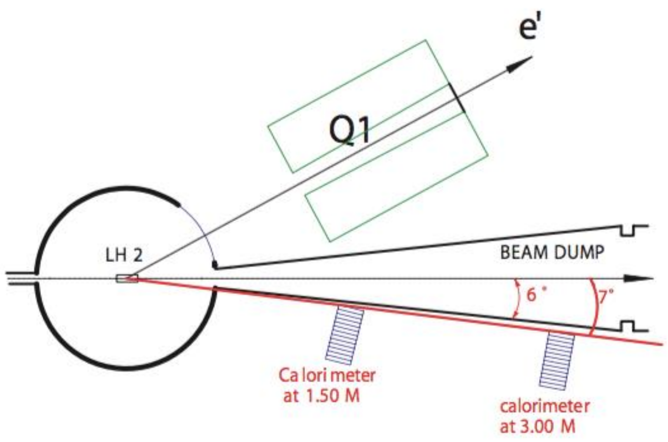


Detection topologies

- A reaction is **exclusive** when the final state is **fully determined**
- Depending on the **detector properties,** on the **final state** one is interested in, the available **statistics** of the data, the possible **competing background channels**, etc, different detection topologies can be adopted

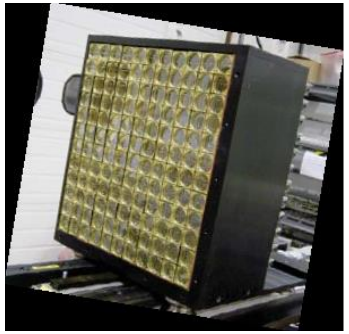
Examples:

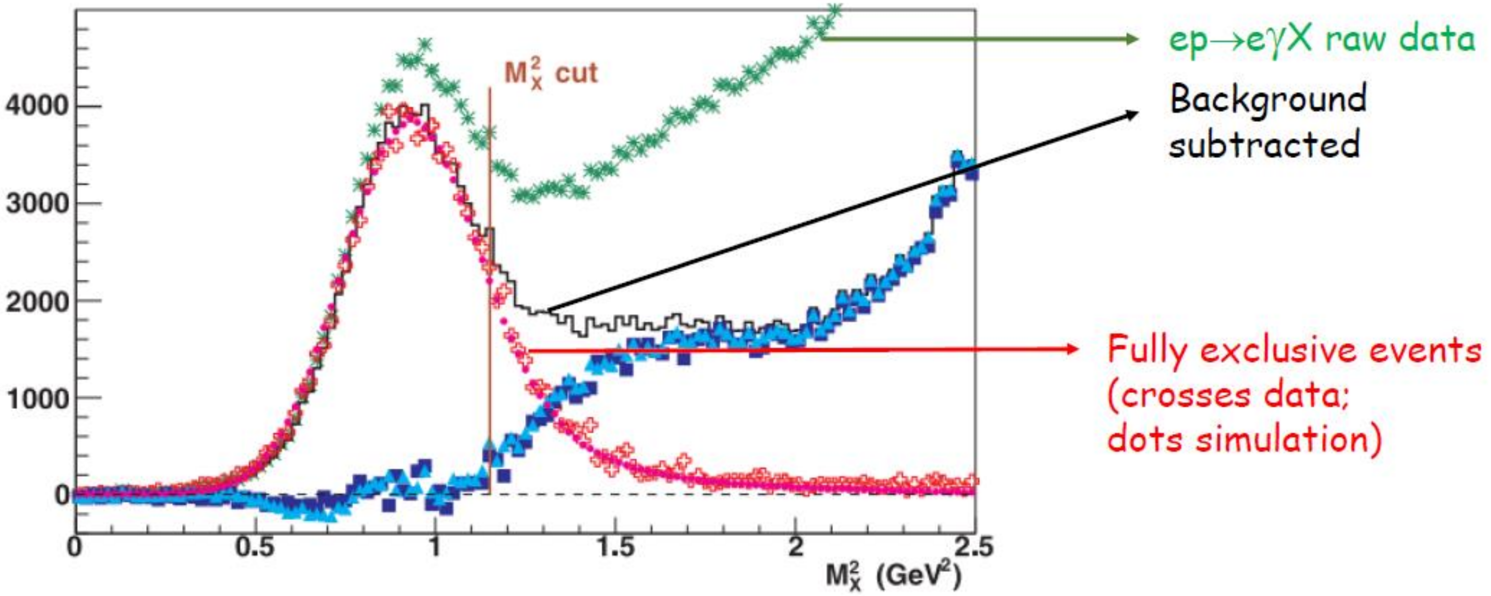
- DVCS in Hall A: eγX topology, the proton is reconstructed via MISSING MASS **ep→e'γX** $M_X^2 = \left(p_e^\mu + p_{target}^\mu - p_{e'}^\mu - p_\gamma^\mu\right)^2$



Electron detected in High-Resolution Spectrometer

Photon detected in dedicated calorimeter



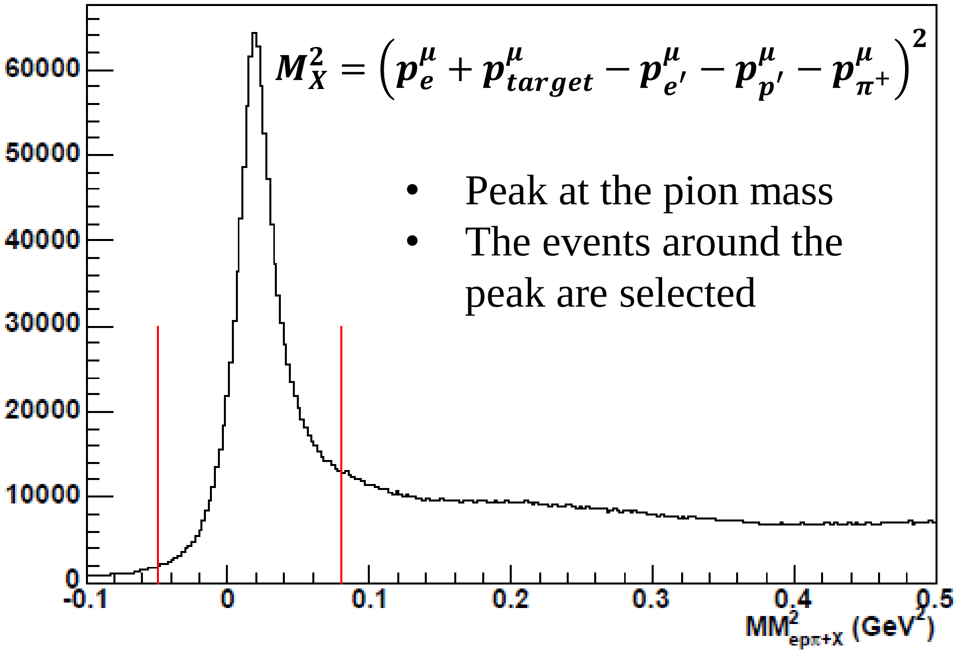


Detection topologies

- A reaction is **exclusive** when the final state is **fully determined**
- Depending on the **detector properties,** on the **final state** one is interested in, the available **statistics** of the data, the possible **competing background channels**, etc, different detection topologies can be adopted

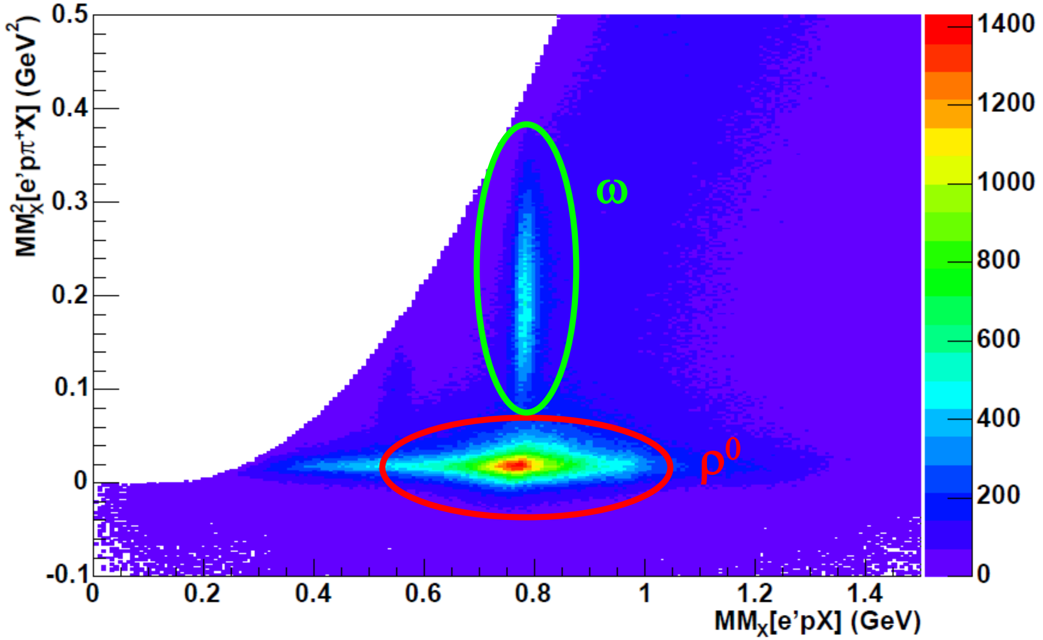
Examples:

- DVCS in Hall A: eγX topology, the proton is reconstructed via MISSING MASS **ep→e'γX** $M_X^2 = \left(p_e^\mu + p_{target}^\mu - p_{e'}^\mu - p_\gamma^\mu\right)^2$
- DVCS and exclusive π^0 in CLAS/CLAS12: epγ and epγγ topology, many *exclusivity variables* are available (next slides)
- DVMP (ρ^0) in CLAS: **ep→e'p'ρ⁰→e'p'π⁺(π⁻)**
  - Higher acceptance for positively charged particles due to the torus magnetic field setting (negatives inbending)
  - Overall higher efficiency when detecting 3 particles instead of 4

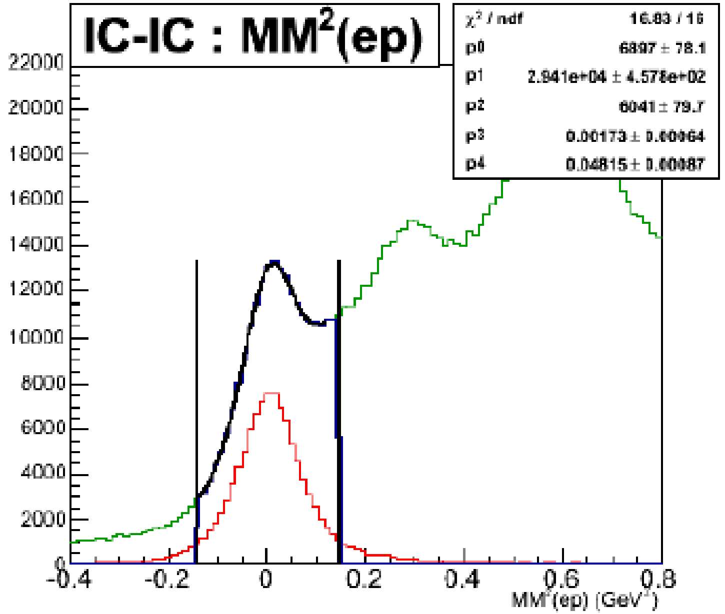


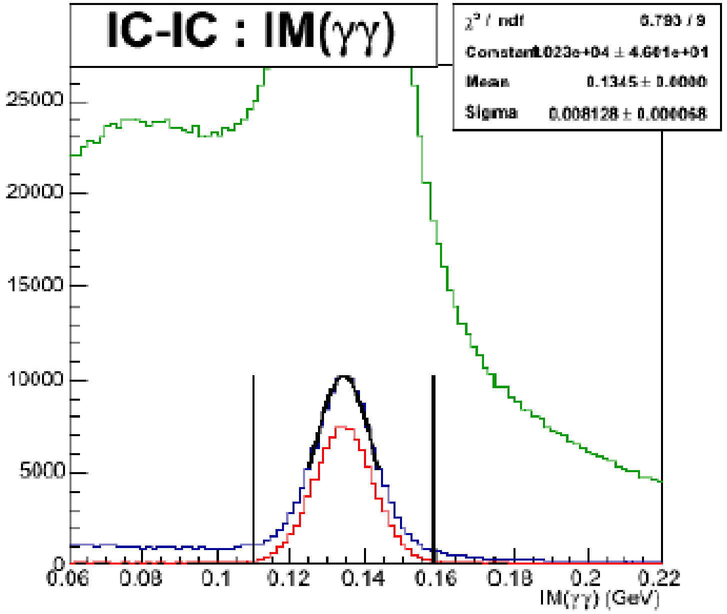
$M_X^2 = \left(p_e^\mu + p_{target}^\mu - p_{e'}^\mu - p_{p'}^\mu - p_{\pi^+}^\mu\right)^2$

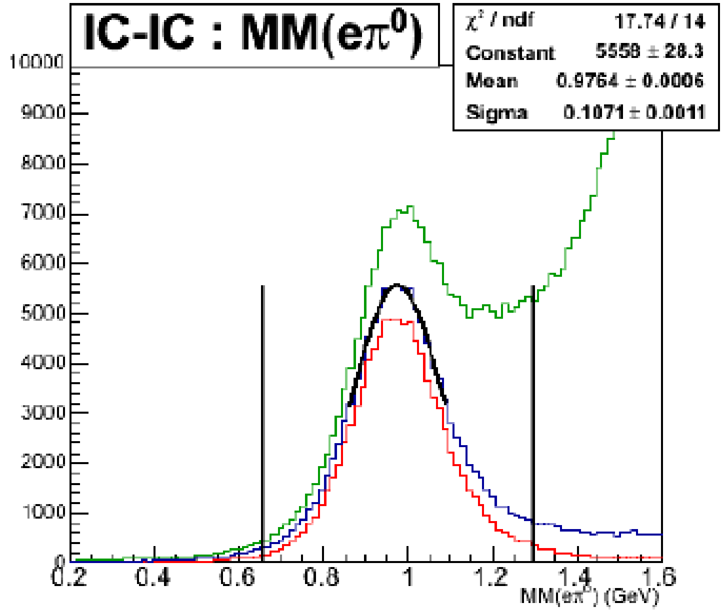
- Peak at the pion mass
- The events around the peak are selected

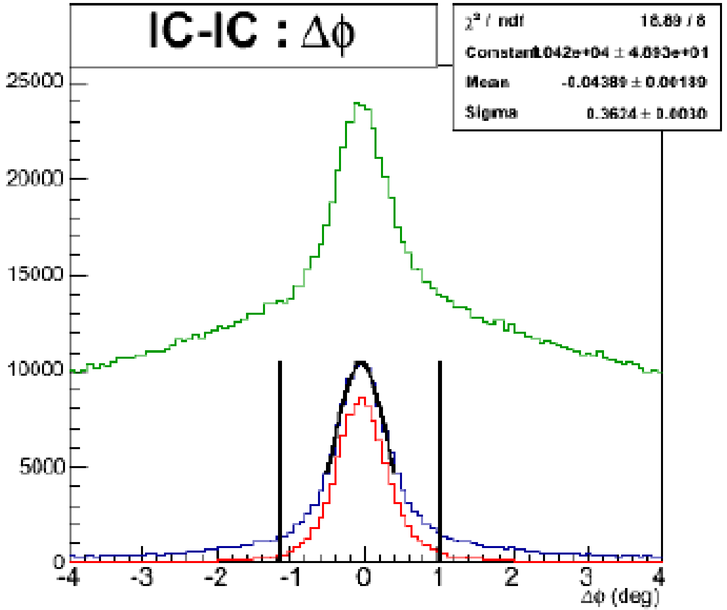


Exclusivity cuts (case ep→e’p’π^0)









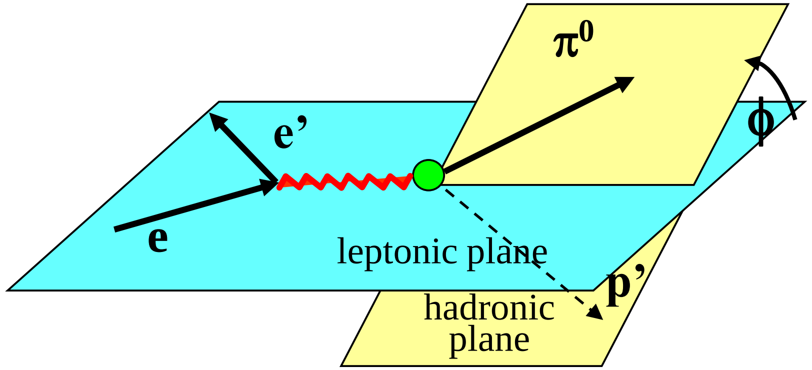
Detection topology: epγγ

Exclusivity variables:

- Two-photons invariant mass
- Missing mass ep→e’p’X
- Missing mass ep→e’π^0X
- Coplanarity (Δϕ): difference between two ways to compute the angle between hadronic and leptonic plane

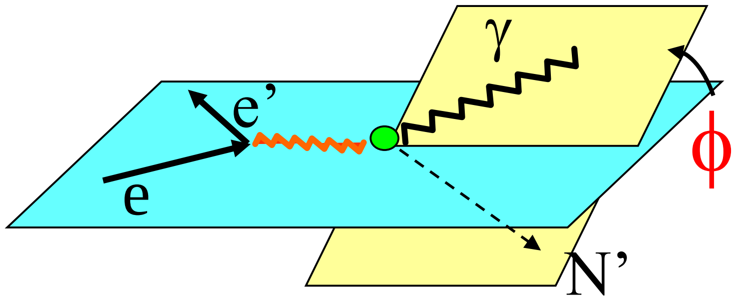
3-σ cuts (gaussian fits)

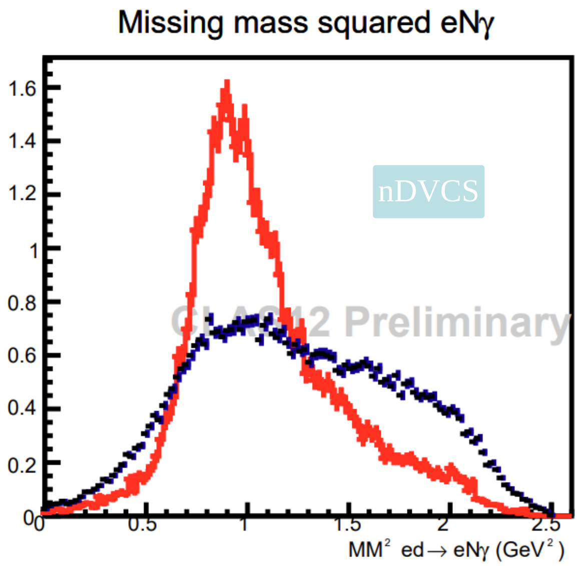
→ very clean final event sample

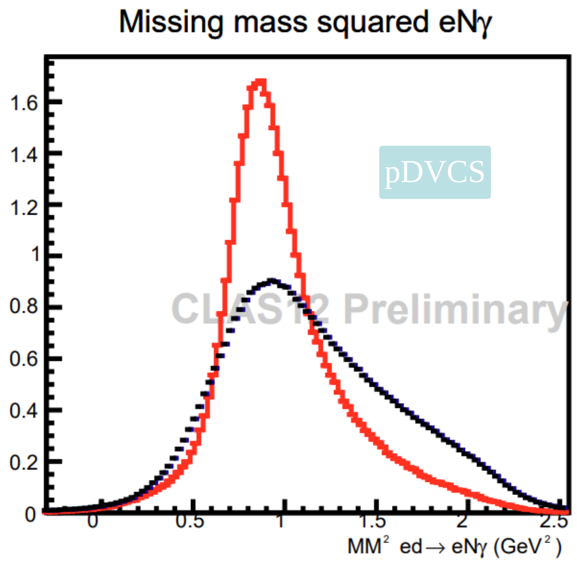


Exclusivity cuts (case of ed→e’nγ(p), e’pγ(n))

- Events with at least one electron, photon, and nucleon (n or p) are selected
- The nDVCS (pDVCS) final state is further selected with the following exclusivity criteria: (N:nucleon)
  - Missing masses
    - e d → e N γ X (must peak at proton mass)
    - e N → e N γ X (must peak at 0)
    - e N → e N X (must peak at photon mass = 0)
  - Missing momentum
    - e d → e N γ X (must give the momentum distribution of the spectator nucleon)
  - Δϕ, Δt, θ(γ,X)
    - Difference between two ways of calculating ϕ and t
    - Cone angle between measured and reconstructed photon X
- Data/MC comparison is used to determine size of cuts




Missing mass squared eNγ


Missing mass squared eNγ

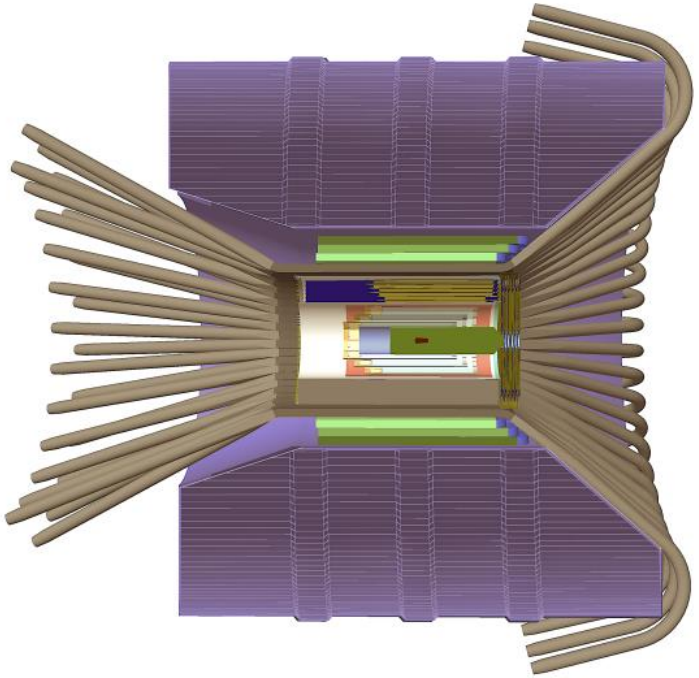
● Simulation

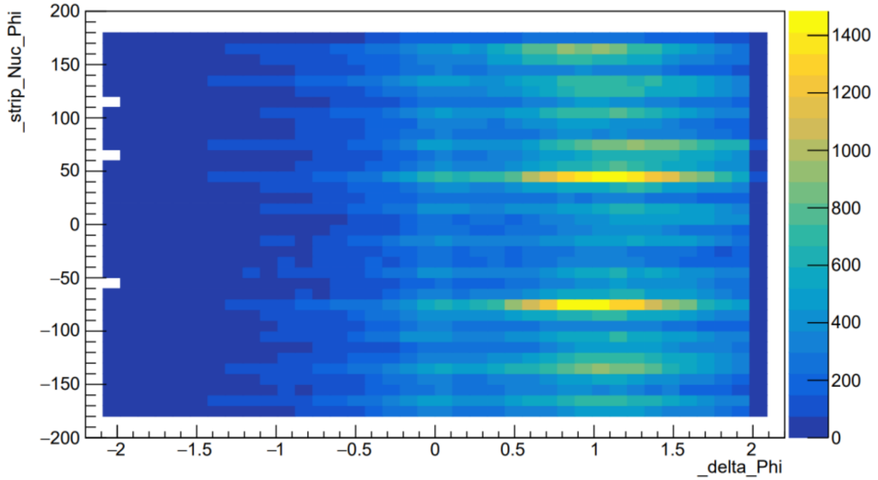
● Data with π^0 contamination

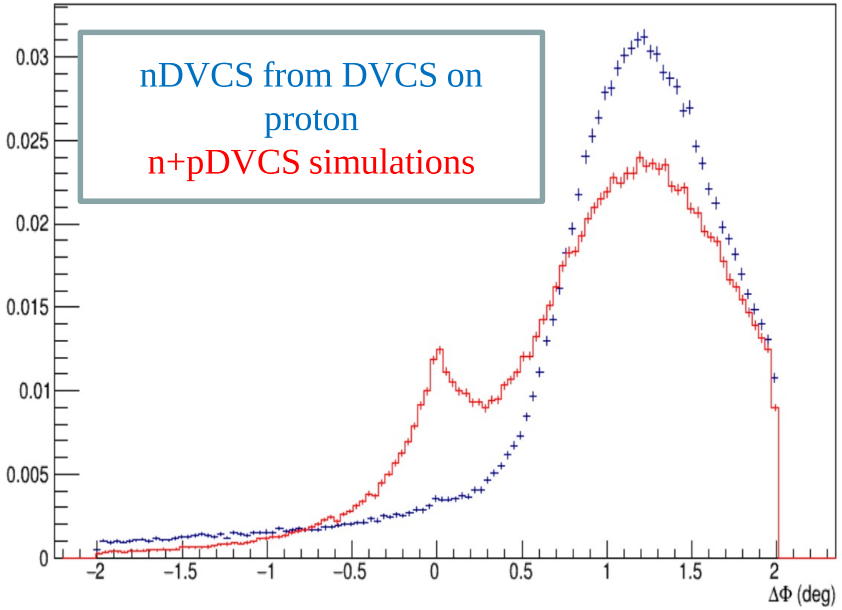
π^0background contamination is estimated using simulations (see one of the next slides)

When standard PID is not enough: improving the neutron selection with ML techniques

- The tracking of the Central Vertex Tracker is not 100% efficient nor uniform
- In the dead regions of the CVT protons have no associated track and thus can be misidentified as neutrons
- Protons roughly account for more than >40% contamination in the “nDVCS” signal sample
- Approach based on Machine Learning & Multi-Variate Algorithms:
  - Reconstruct nDVCS from DVCS experiment on proton requiring neutron PID: selected neutron are misidentified protons
  - Use this sample to determine the characteristics of fake neutrons in low- and high-level reconstructed variables
  - Based on those characteristics, subtract the fake neutrons contamination from nDVCS
  - As a « signal » sample in the training of the ML, use $ep \rightarrow en\pi^+$events from DVCS experiment on proton

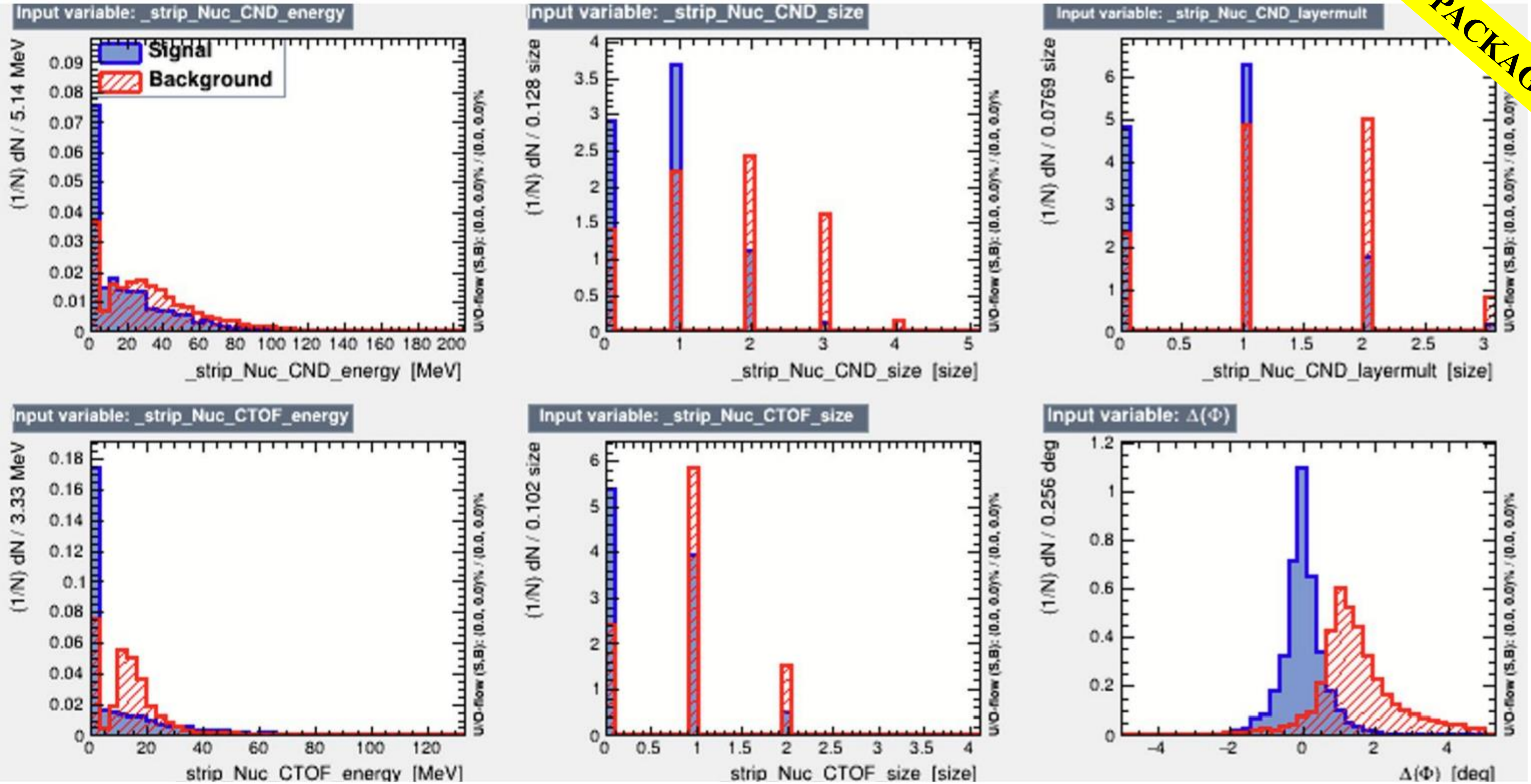






Improving the neutron selection with ML techniques

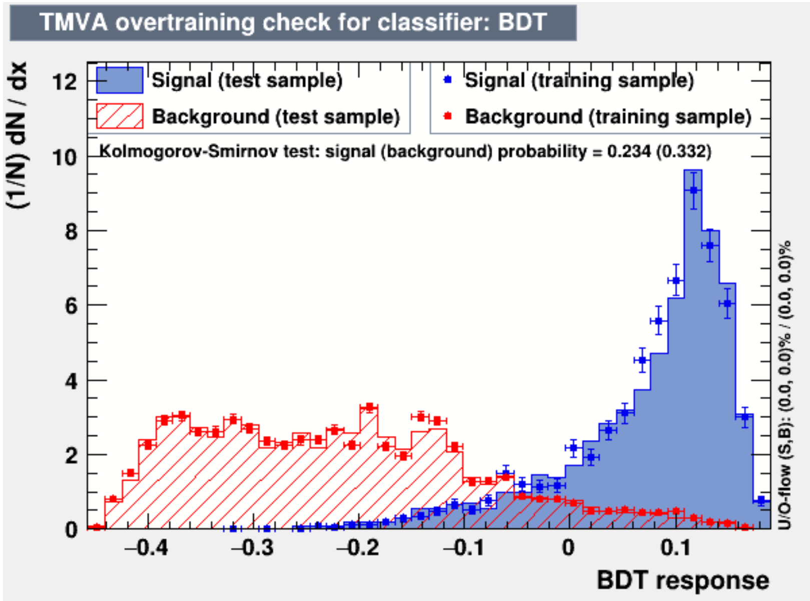
ROOT PACKAGE TMVA

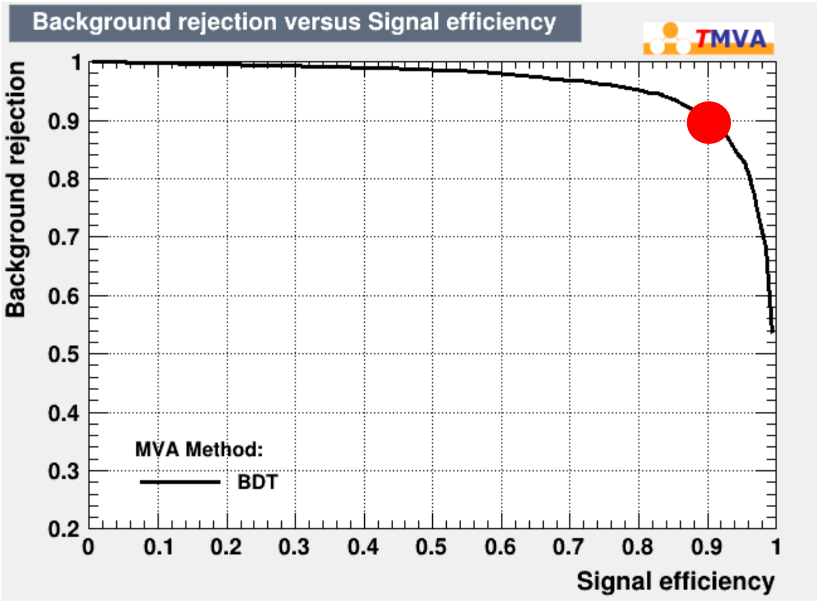


Using detector variables (CTOF and CND) and one exclusivity variable ($\Delta\phi$)

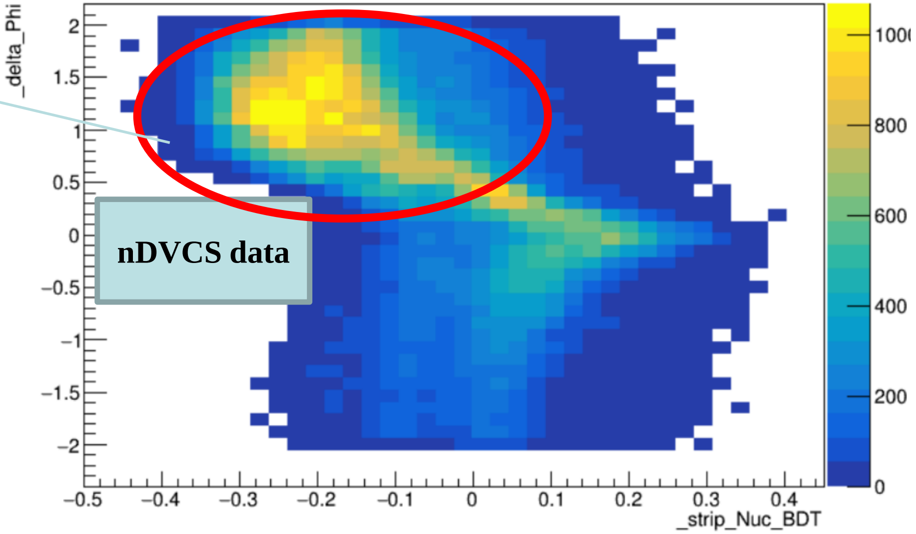
Improving the neutron selection with ML techniques

- Advantages of the ML tool:
- Directly trained on data
- Better optimization of signal to background ratio





Tool isolates contamination

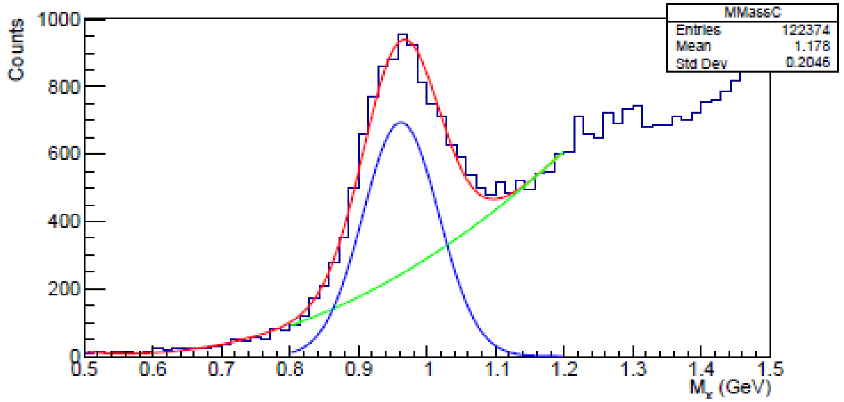


Depending where one places the « cut » on the BDT response, the rejection power and efficiency vary.

- **~90% of background rejection**
- **~90% signal efficiency**

Positron ID with ML (from TCS analysis @ CLAS12)

HTCC: good PID for positrons with p up to 4.9 GeV; for p>4.9 GeV, π^+'s can be misidentified as positrons



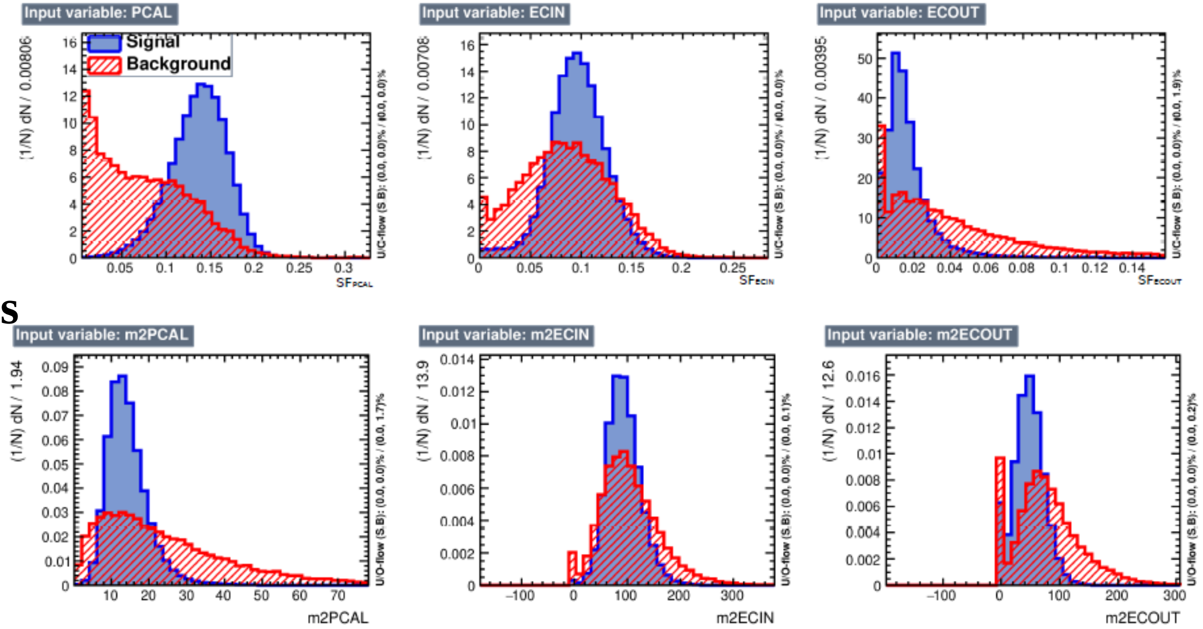
Signature of contamination: ee^+X missing mass peaking at the neutron mass (p>4.4 GeV, MM computed attributing the π^+ mass to the « positrons »):
→ **strong contamination of $e\pi^+(n)$ events**

Signal: e^+ identified as e^+ Background: π^+ identified as e^+

Strategy and discriminating variables

Positron: electromagnetic shower Pion: Minimum Ionizing Particle (MIP)

$$SF_{\mathrm{EC\ Layer}} = \frac{E_{dep}(\mathrm{EC\ Layer})}{P} \qquad M_2 = \frac{1}{3}\sum_{U,V,W} \frac{\sum_{\mathrm{strip}} (x-D)^2 \cdot \ln(E)}{\sum_{\mathrm{strip}} \ln(E)}$$



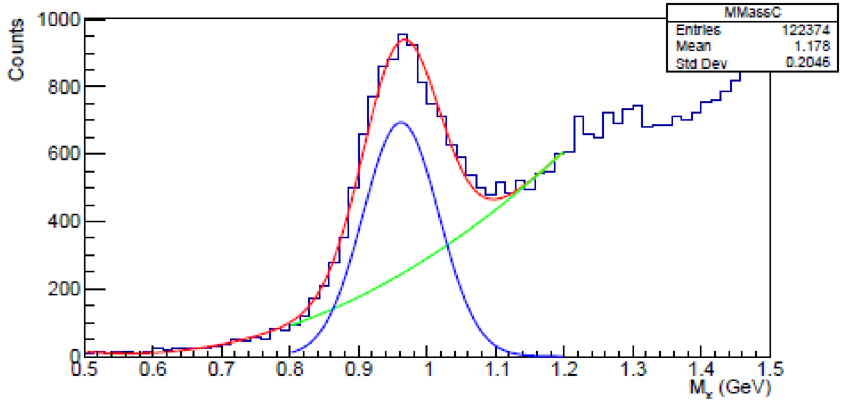
→ 6 variables (from PCAL, Ecin, Ecout): SF and shower widths

- Signal in data ⇒ Outbending electrons
- Background in data ⇒ $ep \to e\pi^+_{PID=e^+}(n)$

Positron ID with ML (from TCS analysis @ CLAS12)

HTCC: good PID for positrons with p up to 4.9 GeV; for p>4.9 GeV, π^+'s can be misidentified as positrons

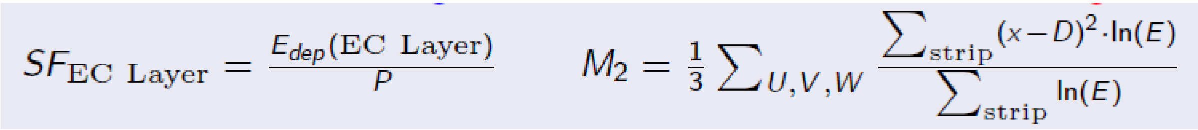
Signature of contamination: ee^+X missing mass peaking at the neutron mass (p>4.4 GeV, MM computed attributing the π^+ mass to the « positrons »):
→ **strong contamination of $e\pi^+(n)$ events**



Signal: e^+ identified as e^+ Background: π^+ identified as e^+

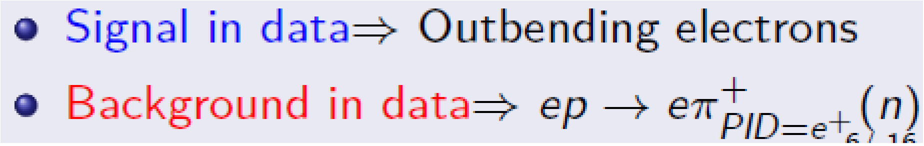
Strategy and discriminating variables

Positron: electromagnetic shower Pion: Minimum Ionizing Particle (MIP)

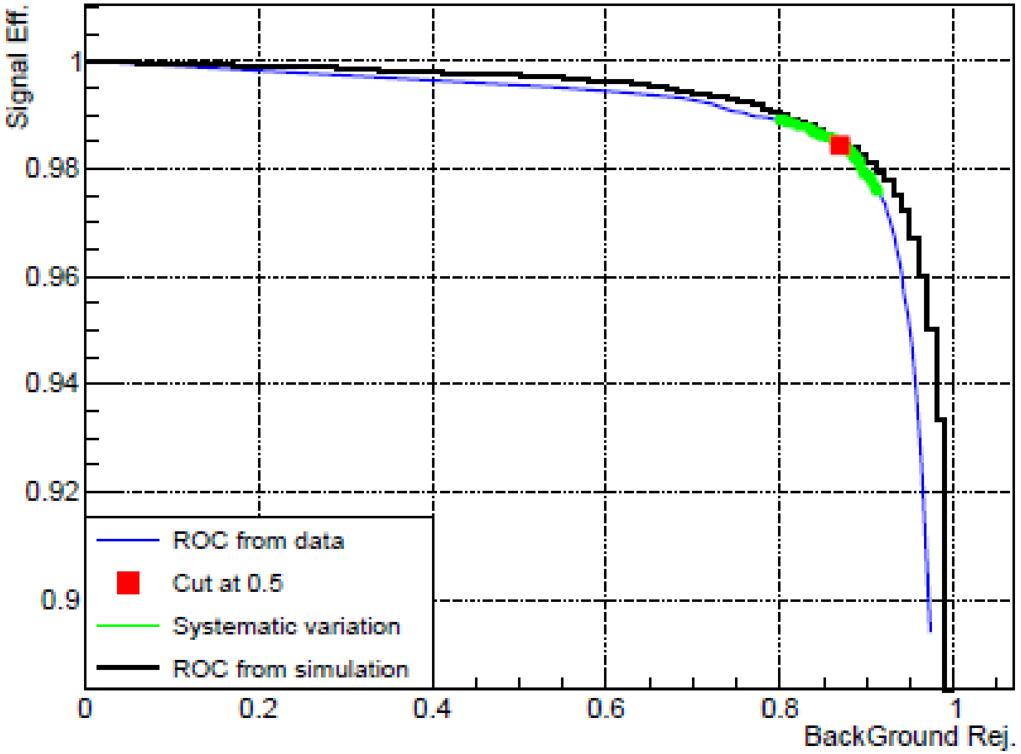


$$SF_{\text{EC Layer}} = \frac{E_{dep}(\text{EC Layer})}{P} \qquad M_2 = \frac{1}{3}\sum_{U,V,W} \frac{\sum_{\text{strip}} (x-D)^2 \cdot \ln(E)}{\sum_{\text{strip}} \ln(E)}$$

→ **6 variables (from PCAL, Ecin, Ecout): SF and shower widths**



- Signal in data ⇒ Outbending electrons
- Background in data ⇒ $ep \rightarrow e\pi^+_{PID=e^+}(n)$

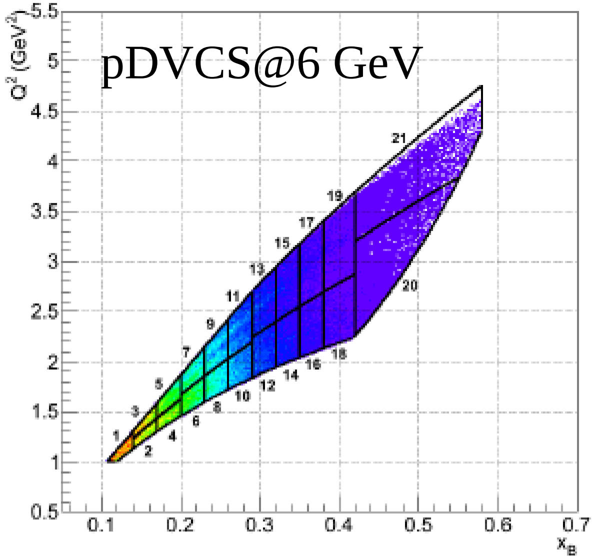


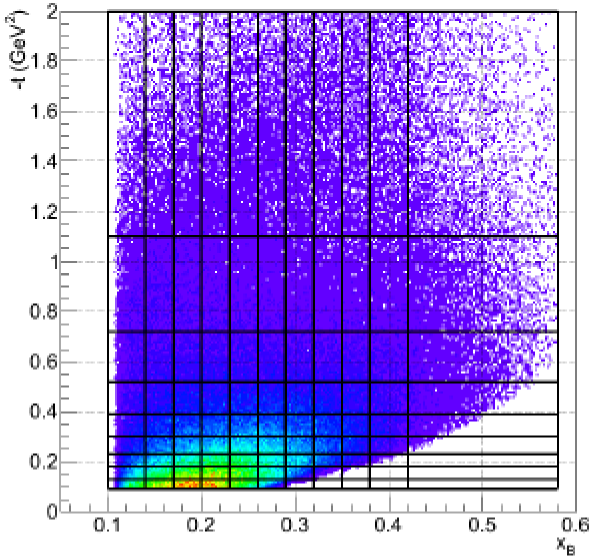
Binning

- CFFs depend on x_B (→ξ), t, and (mildly) Q^2; they are extracted from the ϕ modulations of observables

→ a multi-dimensional and fine binning is necessary to sample the complex kinematic dependence of GPDs

- The higher the statistics (cross section * beam time * acceptance), the finer the binning that can be done
- A fine binning prevents strong variations of the kinematic within the bin → easier comparison with theory
- The binning should never be finer than the experimental resolution on each of the variables that we bin in





nDVCS@11 GeV

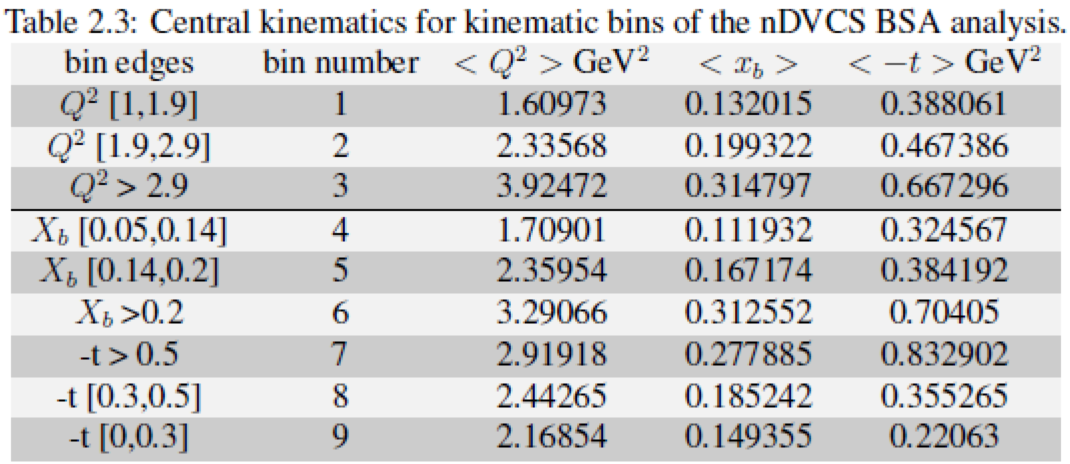


Table 2.3: Central kinematics for kinematic bins of the nDVCS BSA analysis.

bin edges	bin number	$< Q^2 >$ GeV^2	$< x_b >$	$< -t >$ GeV^2
Q^2 [1,1.9]	1	1.60973	0.132015	0.388061
Q^2 [1.9,2.9]	2	2.33568	0.199322	0.467386
$Q^2 > 2.9$	3	3.92472	0.314797	0.667296
X_b [0.05,0.14]	4	1.70901	0.111932	0.324567
X_b [0.14,0.2]	5	2.35954	0.167174	0.384192
X_b >0.2	6	3.29066	0.312552	0.70405
-t > 0.5	7	2.91918	0.277885	0.832902
-t [0.3,0.5]	8	2.44265	0.185242	0.355265
-t [0,0.3]	9	2.16854	0.149355	0.22063

Backgrounds: π^0 subtraction from eNγ events

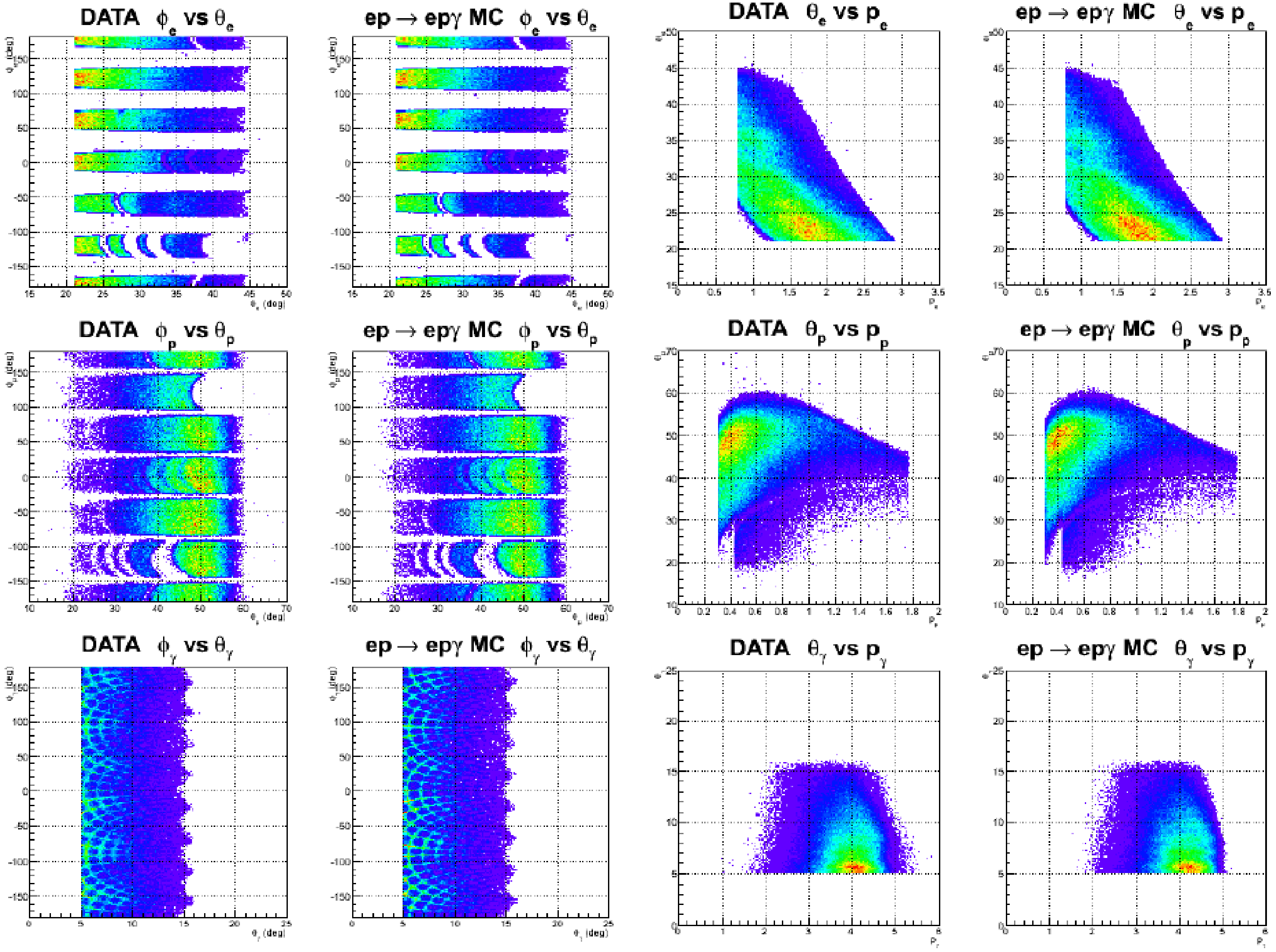
- After exclusivity cuts, the eNγ event sample will still contain contamination from eNπ^0 events for which **one of the two decay photons has escaped detection**
- Subtraction of these π^0 events is done, bin by bin, using simulations of the background channel
- Monte Carlo simulations:
  - GPD-based event generator for π^0
  - Response of the detector simulated by GEANT4 simulation of CLAS12 (GEMC)
- Description of the method:
  - Estimate the ratio of partially reconstructed eN π^0(1 photon) decay to fully reconstructed eN π^0 decays in MC
  - This is done for each kinematic bin to minimize MC model dependence
  - Multiply this ratio by the number of reconstructed eN π^0 in data to get the number of eN π^0(1 photon) in data
  - Subtract this number from DVCS reconstructed decays in data for each kinematical bin

Simulations: $R = N(eN\pi^0_{1\gamma})/N_{MC}(eN\pi^0)$

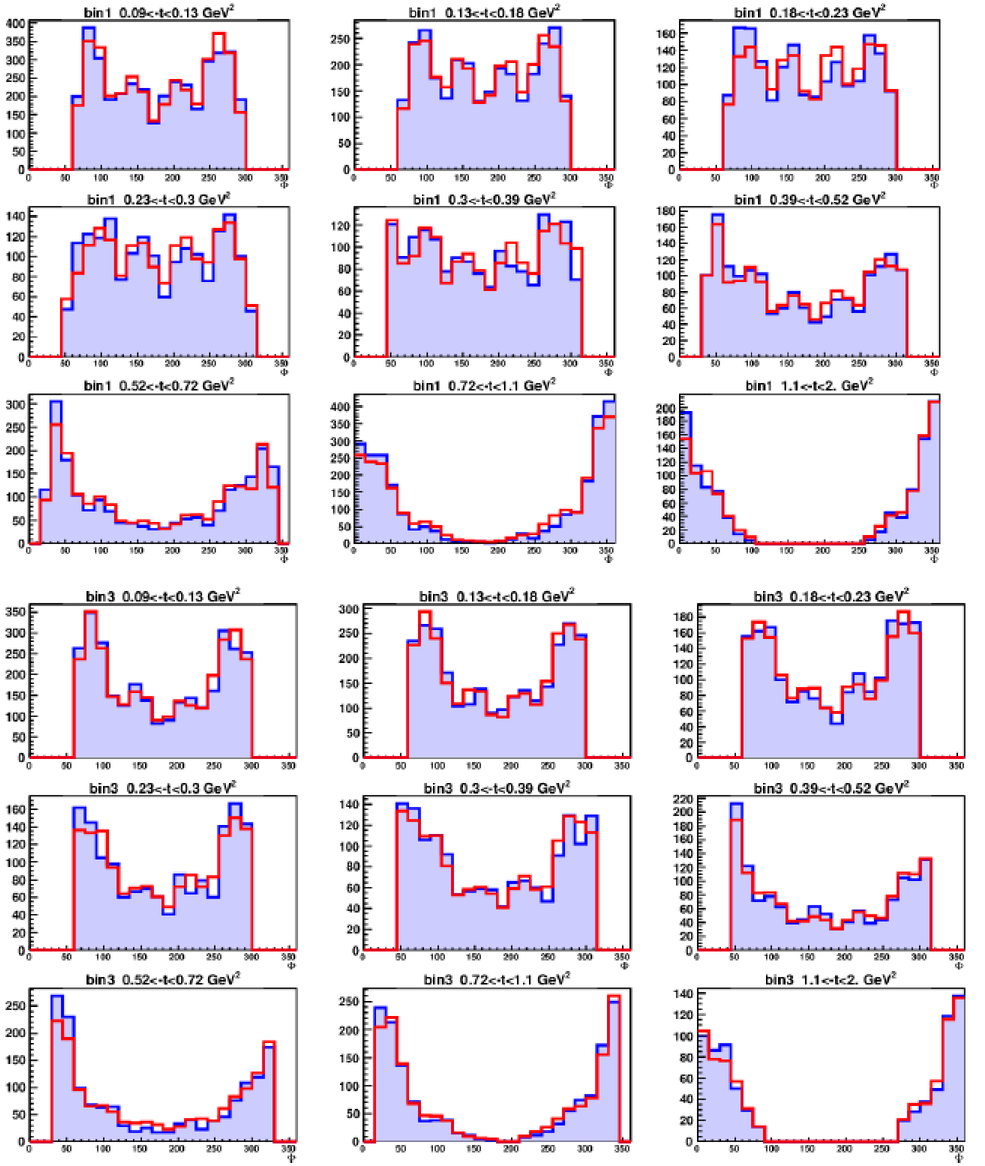
Data + simulation: $N\left(eN\pi^0_{1\gamma}\right) = R * N_{DATA}(eN\pi^0) \rightarrow \boldsymbol{N(DVCS) = N(DVCS_{recon}) - N(eN\pi^0_{1\gamma})}$

Acceptance*efficiency (for cross section measurements)

- The acceptance*efficiency is computed, for each 4-dimensional kinematic bin, using a Monte-Carlo simulation reproducing the experimental distributions as faithfully as possible (fiducial cuts)



MC
Data

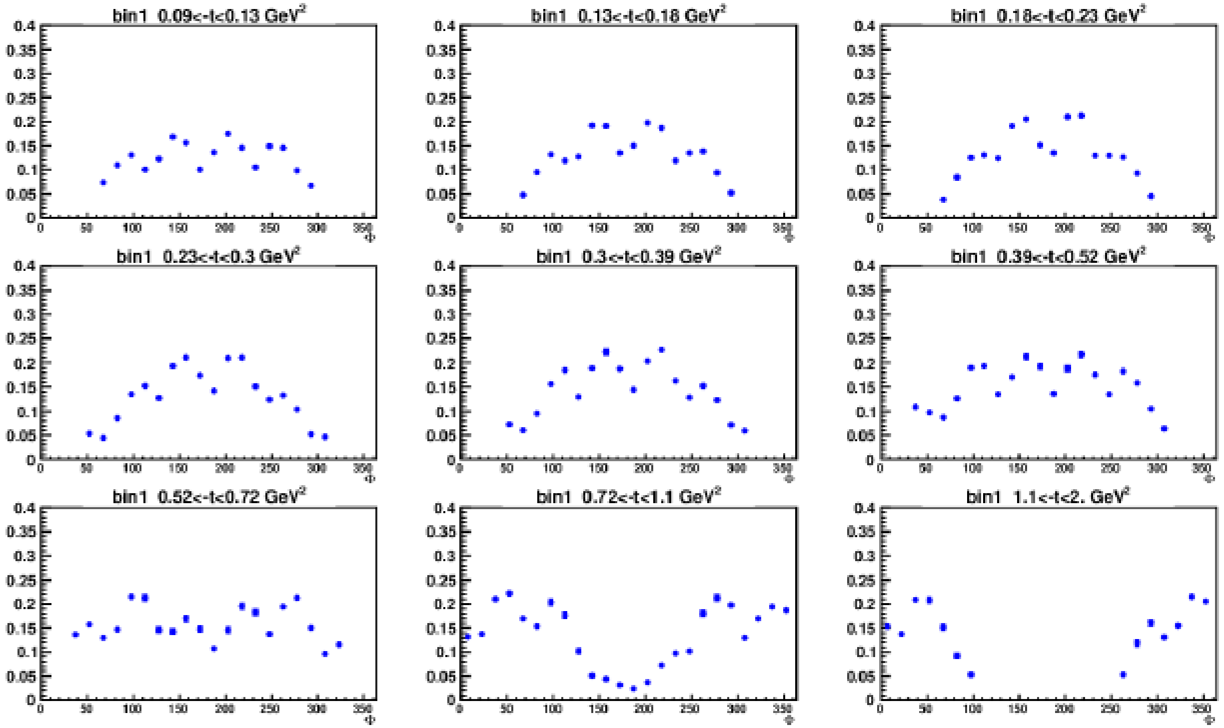


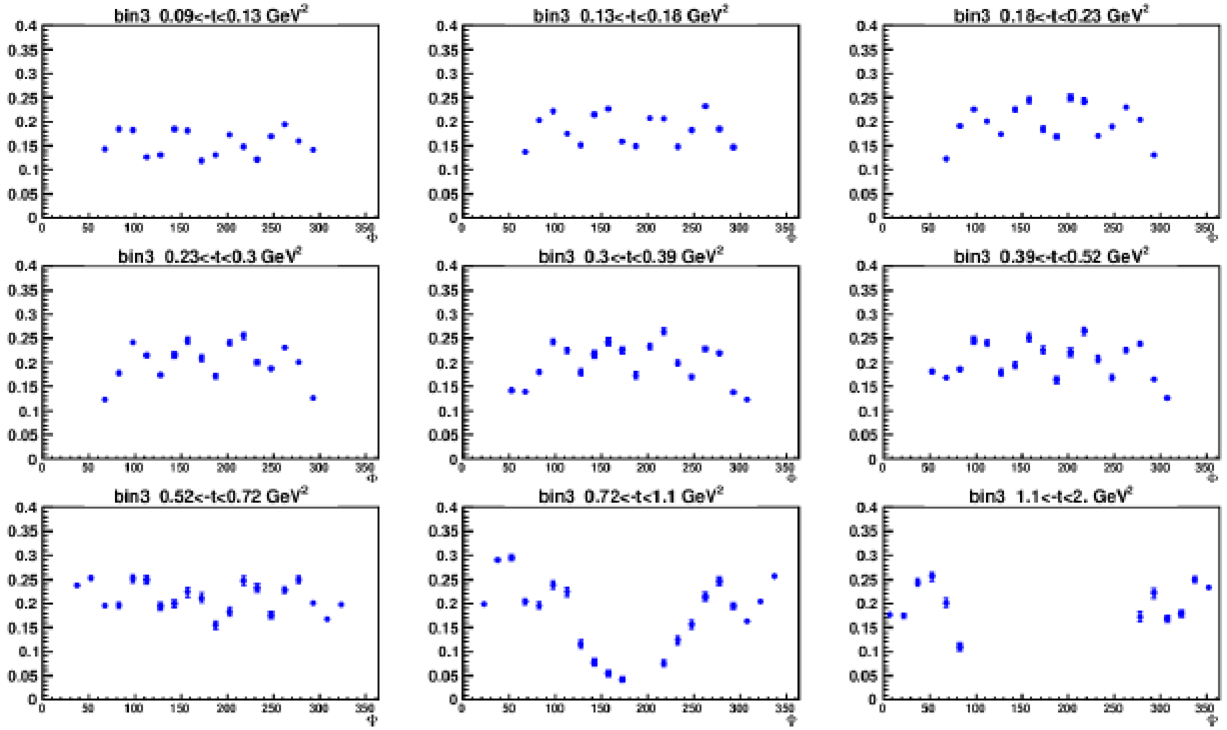
Acceptance/efficiency (for cross section measurements)

- The acceptance*efficiency is computed, for each 4-dimensional kinematic bin, using a Monte-Carlo simulation reproducing the experimental distributions as faithfully as possible (fiducial cuts)
- Definition:

$$Acc(Q^2, x_B, t, \phi) = \frac{N_{rec}(Q^2, x_B, t, \phi)}{N_{gen}(Q^2, x_B, t, \phi)}$$

N_{gen} : number of $ep \to ep\gamma$ MC events generated in the (Q^2, x_B, t, ϕ) bin
N_{rec} : number of events reconstructed in the bin

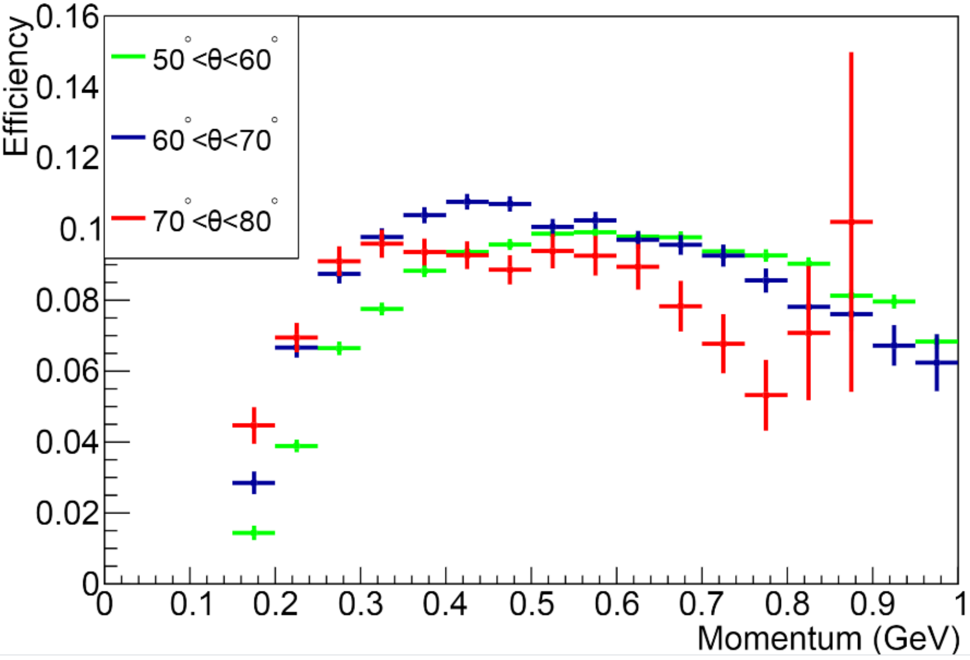




When Monte-Carlo is not enough: data-driven efficiencies

After the CND was installed and commissioned, its efficiency for neutrons was verified using beam data, with hydrogen target

Neutron efficiency from ep→e'nπ^+ (RGA data)



P. Chatagnon et al., NIM A 959 (2020)

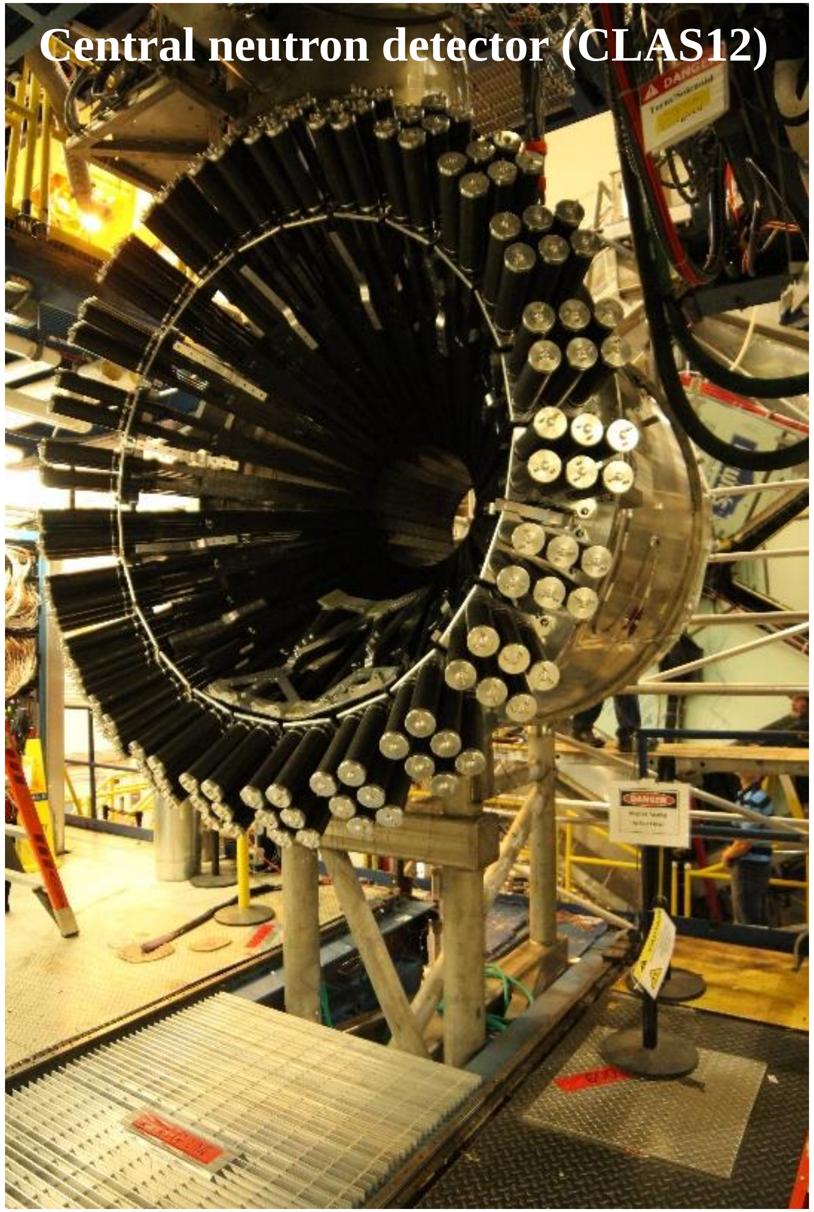
The ep→e'nπ^+ final state is chosen because:

- it is an exclusive final state, fully determined, with a neutron in it
- it can be reconstructed either detecting all particles or by missing mass

Detected in CND

$$Eff = \frac{\#detected\ en\pi^+}{\#MM\ e\pi^+X}$$

Cut around the neutron mass



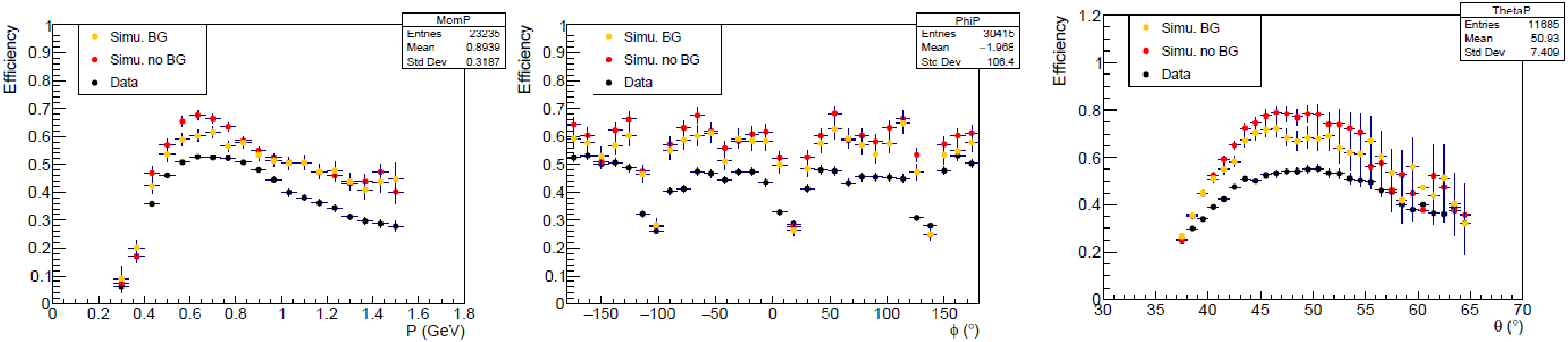
Central neutron detector (CLAS12)

When Monte-Carlo is not enough: data-driven efficiencies

Proton efficiency in the Central Detector

ep → e(p)ρ^0→e(p)$\pi^+\pi^-$ with 0<6 GeV < $M_{\pi+\pi-}$ < 1 GeV

$$Eff = \frac{\#detected\ ep\pi^+\pi^-}{\#MM\ e\pi^+\pi^-X}$$



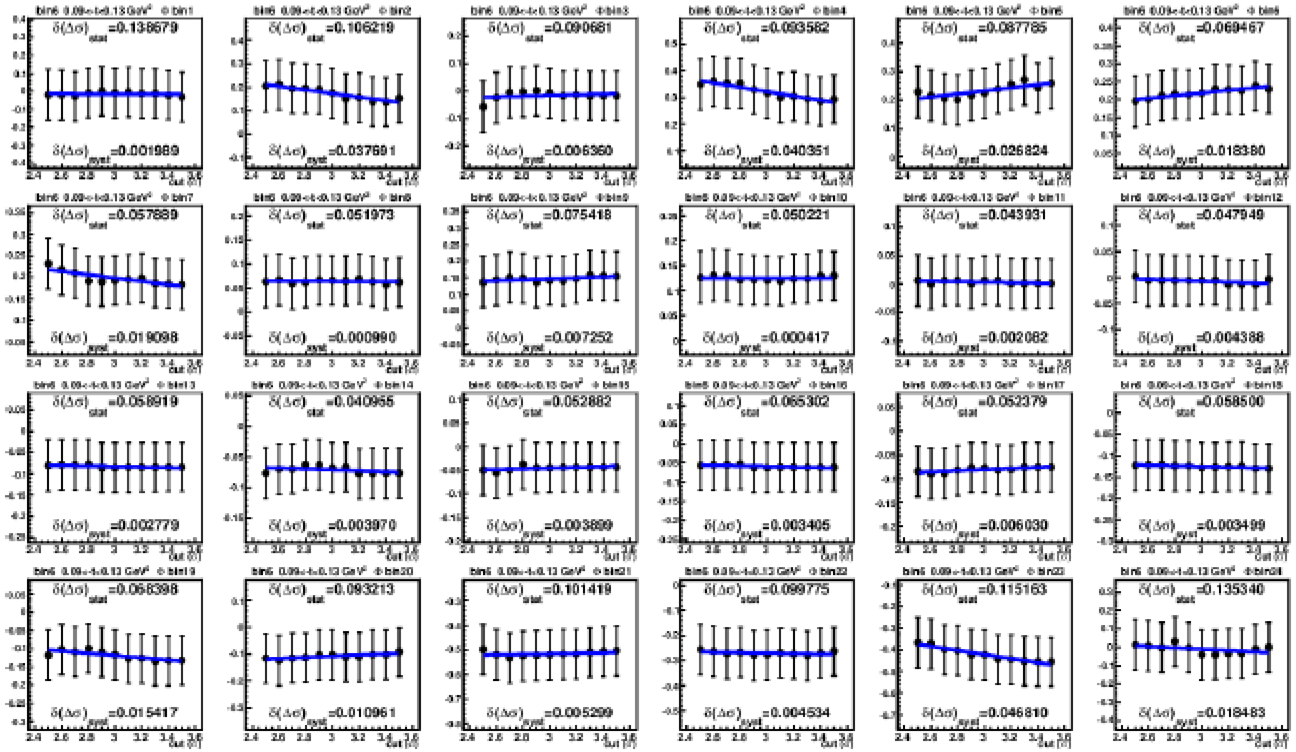
Correction to the Monte-Carlo efficiency ~20%

AKA: another cool application of exclusive reactions ☺
A useful tool to verify detector performance

Systematic uncertainties

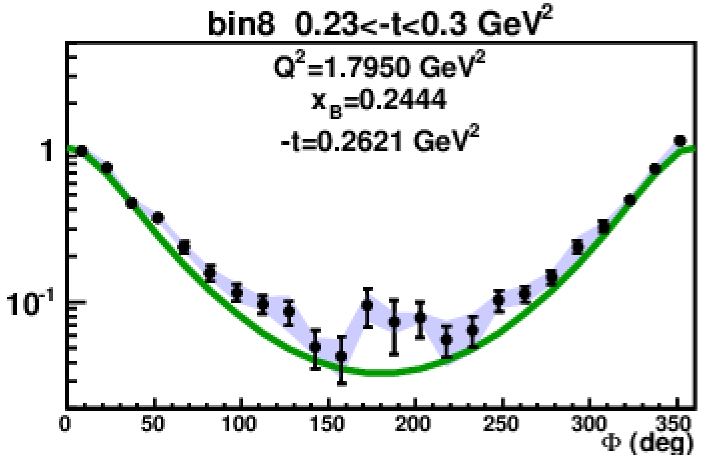
Systematic uncertainties (from Google ☺): ***« a possible unknown variation in a measurement, or in a quantity derived from a set of measurements, that does not randomly vary from data point to data point."***

Each choice we make in an analysis can be a source of systematic uncertainties: PID cuts, model chosen for the Monte-Carlo event generator, exclusivity cuts, etc…
Systematics are computed trying different choices and comparing the obtained results

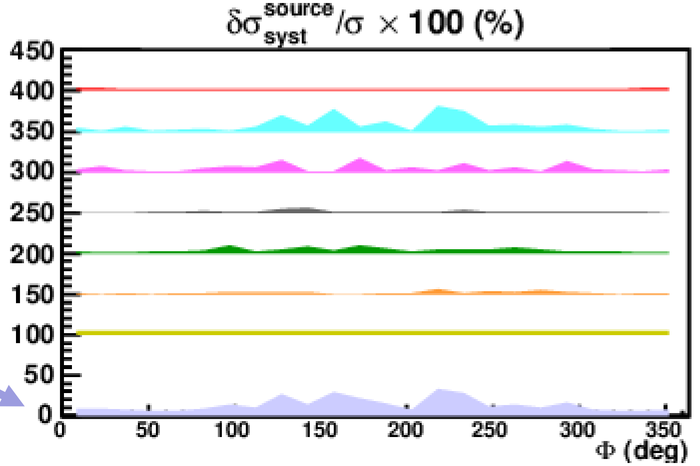


For each bin $\delta\theta_{syst}$ *(exclusivity cuts) = 1/2 |f(3.5) − f(2.5)|*
where *f*: fitted function

Computing systematics means redoing the whole analysis over and over… ☹



The systematic uncertainties obtained from all sources are typically shown separately and/or combined in a quadratic sum



Summary

Exclusive reactions are a unique tool to explore the multi-dimensional structure of the nucleon via the extraction of Generalized Parton Distributions

Measuring cross sections and asymmetries for exclusive reactions is challenging for several reasons:

- Small cross sections
- Need to extract observables as a function of several kinematic variables
- Multiparticle final states → limited detection efficiency
- Various kinds of backgrounds

Multi-purpose, large acceptance detectors such as CLAS12 are ideal setups for these kinds of measurements

Various analysis techniques can be employed to extract the observables → creativity is the key ☺

Thank you all for attending my lectures, it has been a pleasure and a honor.
Best of luck for your PhD's, your lives and your careers!